Asymptotic properties of the maximum likelihood estimator of random effects models with serial correlation

Jimmy Skoglund[*]and Sune Karlsson[†]
Stockholm School of Economics
Department of Economic Statistics

Stockholm School of Economics
SSE/EFI Working Paper Series in Economics and Finance
No. 432

February 2001

Abstract

This paper considers the large sample behavior of the maximum likelihood estimator of random effects models with serial correlation in the form of AR(1) for the idiosyncratic or time-specific error component. Consistent estimation and asymptotic normality as N and/or T grows large is established for a comprehensive specification which nests these models as well as all commonly used random effects models. When N or $T \to \infty$ only a subset of the parameters are consistent and asymptotic normality is established for the consistent subsets.

Keywords: Panel data; serial correlation; random effects.

JEL codes: C12; C13; C23

[*]Jimmy.Skoglund@hhs.se

[†]Sune.Karlsson@hhs.se

1 Introduction

Ever since the seminal work of Balestra and Nerlove (1966) there has been a large interest in and use of random effects models. An important further development was the generalization of the one-way model with individual effects to allow for serial correlation by Lillard and Willis (1978). This model captures correlation in the data at the individual level and has been elaborated by, among others, Anderson and Hsiao (1982), MaCurdy (1982) and Baltagi and Li (1991, 1994). This is, however, not the only conceivable source of correlation. It is quite reasonable to expect random time effects to be correlated as well — reflecting serial correlation in the variables driving unobserved time specific heterogeneity. There are, consequently, a number of variations on random effects models allowing for correlation in the time effect. King (1986) studies a one-way model with serially correlated time effects, Magnus and Woodland (1988) consider a one-way model with both serially correlated time effects and idiosyncratic errors in a multivariate setting and Revankar (1979) proposed a two-way model with serially correlated time effects. Recently Karlsson and Skoglund (2000) derived a straightforward maximum likelihood estimator as well as hypothesis tests for the latter model.

While random effects models with serial correlation in the error components are being used extensively in empirical work the theoretical aspects are less well developed. Anderson and Hsiao (1982) consider the consistency properties of the one-way model with individual effects and serially correlated idiosyncratic effects. Amemiya (1971) proves the consistency and asymptotic normality of the maximum likelihood estimator of the standard two-way model as both N and T grows large. This paper extends the work of Anderson and Hsiao and Amemiya by establishing the asymptotic properties of a comprehensive random effects specification which nests the one-way models with serial correlation as well as the two-way model with serial correlation. More specifically, the model of interest is

$$\begin{aligned} y_{it} &= \alpha + \mathbf{x}_{it}'\boldsymbol{\beta} + \mathbf{d}_t'\boldsymbol{\pi} + \mathbf{h}_i'\boldsymbol{\tau} + \varepsilon_{it} \\ \varepsilon_{it} &= \mu_i + \lambda_t + v_{it} \end{aligned} \tag{1}$$

with λ_t an AR(1),

$$\lambda_t = \rho_\lambda \lambda_{t-1} + u_t, \tag{2}$$

and v_{it} an AR(1),

$$v_{it} = \rho_v v_{it-1} + e_{it}, \tag{3}$$

where $\mathbf{x}_{it}$ varies over both individuals and time, $\mathbf{d}_t$ is individual-invariant and $\mathbf{h}_i$ is time-invariant. If there are no time effects we obtain the one-way model with individual effects and serially correlated idiosyncratic errors and if there are no individual effects we obtain the one-way model with both serially correlated time effects and serially correlated idiosyncratic errors. Setting $\rho_v = 0$ obtains the two-way model with serially correlated time effects and setting $\rho_\lambda = 0$ obtains a model not discussed previously in the literature. That is, the two-way model with serially correlated idiosyncratic errors and independent time effects. The standard one-way models and the standard two-way model are, of course, nested in this specification as well.

In contrast to the earlier literature we consider both consistency and asymptotic normality with traditional large N and fixed T as well as with large T fixed N and both N and T large. We also pay special attention to the effects of including time or individual-invariant explanatory variables in the model.

The organization of the paper is as follows. Section 2 presents the comprehensive specification and the corresponding maximum likelihood estimator. Section 3 derives the asymptotic properties and section 4 concludes with some final remarks. All the proofs are in appendix B.

2 The comprehensive specification

In matrix form the comprehensive model is written

$$\begin{aligned}\mathbf{y} &= \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &= \mathbf{Z}_\mu \boldsymbol{\mu} + \mathbf{Z}_\lambda \boldsymbol{\lambda} + \boldsymbol{\nu}\end{aligned}$$

with $\mathbf{Z}_\mu = (\mathbf{I}_N \otimes \boldsymbol{\iota}_T)$, $\mathbf{Z}_\lambda = (\boldsymbol{\iota}_N \otimes \mathbf{I}_T)$, $\mathbf{Z} = (\boldsymbol{\iota}_{NT}, \mathbf{X}, \mathbf{D}, \mathbf{H})$, where $\mathbf{X}$ is k_1-dimensional, $\mathbf{D} = (\boldsymbol{\iota}_N \otimes \mathbf{d}), \mathbf{d} = (\mathbf{d}_1, \ldots, \mathbf{d}_T)'$ is k_2-dimensional and $\mathbf{H} = (\mathbf{h} \otimes \boldsymbol{\iota}_T), \mathbf{h} = (\mathbf{h}_1, \ldots, \mathbf{h}_N)'$ is k_3-dimensional, $k = \sum_{i=1}^3 k_i$, $\boldsymbol{\delta} = (\alpha, \boldsymbol{\beta}', \boldsymbol{\pi}', \boldsymbol{\tau}')'$, $\boldsymbol{\mu}' = (\mu_1, \ldots., \mu_N)$, $\boldsymbol{\lambda}' = (\lambda_1, \ldots., \lambda_T)$ and $\boldsymbol{\iota}_N$ is a vector of ones of dimension N. Throughout we will maintain the assumption that $e_{it} \sim N(0, \sigma_e^2)$, $\mu_i \sim N(0, \sigma_\mu^2)$, $u_t \sim N(0, \sigma_u^2)$ independent of each other and $\mathbf{X}$, $\mathbf{d}$ and $\mathbf{h}$. In addition we assume that $\rho_\lambda, \rho_v \in (-1, 1)$.

The covariance matrix of the combined error term is given by

$$\begin{aligned}\boldsymbol{\Sigma} &= E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \mathbf{Z}_\mu E(\boldsymbol{\mu}\boldsymbol{\mu}')\mathbf{Z}_\mu' + \mathbf{Z}_\lambda E(\boldsymbol{\lambda}\boldsymbol{\lambda}')\mathbf{Z}_\lambda' + E(\boldsymbol{\nu}\boldsymbol{\nu}') \\ &= \sigma_\mu^2(\mathbf{I}_N \otimes \mathbf{J}_T) + \sigma_u^2(\mathbf{J}_N \otimes \boldsymbol{\Psi}_\lambda) + \sigma_e^2(\mathbf{I}_N \otimes \boldsymbol{\Psi}_v)\end{aligned} \tag{4}$$

where $\mathbf{J}_T = \boldsymbol{\iota}_T \boldsymbol{\iota}_T'$ a $T \times T$ matrix of ones and $\sigma_u^2 \boldsymbol{\Psi}_\lambda$ is the covariance matrix of λ and $\sigma_e^2 \boldsymbol{\Psi}_v$ is covariance matrix of v.

Let $\mathbf{A}$ be the covariance matrix of the one-way model with individual specific effects and serially correlated v_{it}. We can then write

$$\boldsymbol{\Sigma} = \mathbf{A} + \sigma_u^2(\boldsymbol{\iota}_N \otimes \mathbf{I}_T)\boldsymbol{\Psi}_\lambda(\boldsymbol{\iota}_N' \otimes \mathbf{I}_T)$$

where

$$\mathbf{A} = \sigma_\mu^2(\mathbf{I}_N \otimes \mathbf{J}_T) + \sigma_e^2(\mathbf{I}_N \otimes \mathbf{\Psi}_v) = \mathbf{I}_N \otimes \left(\sigma_\mu^2 \mathbf{J}_T + \sigma_e^2 \mathbf{\Psi}_v\right)$$

Following Baltagi and Li (1991) let $\mathbf{C}$ be the Prais-Winsten transformation matrix for $\mathbf{\Psi}_v$ and write

$$\begin{aligned}
\mathbf{C}^{-1}\mathbf{C}\left(\sigma_\mu^2 \mathbf{J}_T + \sigma_e^2 \mathbf{\Psi}_v\right)\mathbf{C}'\mathbf{C}^{-T} &= \mathbf{C}^{-1}\left(\sigma_\mu^2 \left(\mathbf{C}\boldsymbol{\iota}_T\right)\left(\mathbf{C}\boldsymbol{\iota}_T\right)' + \sigma_e^2 \mathbf{I}_T\right)\mathbf{C}^{-T} \\
&= \mathbf{C}^{-1}\left(\sigma_\alpha^2 \overline{\mathbf{J}}_T^\alpha + \sigma_e^2 \overline{\mathbf{E}}_T^\alpha\right)\mathbf{C}^{-T}
\end{aligned}$$

where $\sigma_\alpha^2 = d^2\sigma_\mu^2\left(1-\rho_v\right)^2 + \sigma_e^2$, $\overline{\mathbf{J}}_T^\alpha = \boldsymbol{\iota}_T^\alpha \boldsymbol{\iota}_T^{\alpha\prime}/d^2$, $\boldsymbol{\iota}_T^{\alpha\prime} = \left(\alpha, \boldsymbol{\iota}_{T-1}'\right) = \left(\mathbf{C}\boldsymbol{\iota}_T\right)'$ and $\overline{\mathbf{E}}_T^\alpha = \mathbf{I}_T - \overline{\mathbf{J}}_T^\alpha$ with $d^2 = \boldsymbol{\iota}_T^{\alpha\prime}\boldsymbol{\iota}_T^\alpha = \alpha^2 + (T-1)$, $\alpha = \sqrt{\left(1+\rho_v\right)/\left(1-\rho_v\right)}$. We then have

$$\mathbf{A}^{-1} = \mathbf{I}_N \otimes \mathbf{C}'\left(\sigma_\alpha^{-2}\overline{\mathbf{J}}_T^\alpha + \sigma_e^{-2}\overline{\mathbf{E}}_T^\alpha\right)\mathbf{C} = \mathbf{I}_N \otimes \mathbf{A}^*$$

As in Karlsson and Skoglund (2000) we can then write

$$\begin{aligned}
\mathbf{\Sigma}^{-1} &= \mathbf{A}^{-1} - \mathbf{A}^{-1}(\boldsymbol{\iota}_N \otimes \mathbf{I}_T)[\sigma_u^{-2}\mathbf{\Psi}_\lambda^{-1} + N\mathbf{A}^*]^{-1}(\boldsymbol{\iota}_N' \otimes \mathbf{I}_T)\mathbf{A}^{-1} \\
&= \mathbf{I}_N \otimes \mathbf{A}^* - (\boldsymbol{\iota}_N \otimes \mathbf{A}^*)[\sigma_u^{-2}\mathbf{\Psi}_\lambda^{-1} + N\mathbf{A}^*]^{-1}(\boldsymbol{\iota}_N' \otimes \mathbf{A}^*) \\
&= \mathbf{I}_N \otimes \mathbf{A}^* - \sigma_u^2(\boldsymbol{\iota}_N \otimes \mathbf{A}^*)\left[\mathbf{I}_T + N\sigma_u^2\mathbf{\Psi}_\lambda\mathbf{A}^*\right]^{-1}\mathbf{\Psi}_\lambda(\boldsymbol{\iota}_N' \otimes \mathbf{A}^*)
\end{aligned}$$

and

$$|\mathbf{\Sigma}| = |\mathbf{A}^*|^{-N}\left|\mathbf{I}_T + N\sigma_u^2\mathbf{\Psi}_\lambda\mathbf{A}^*\right|$$

Which gives the log-likelihood as

$$\begin{aligned}
l(\boldsymbol{\delta}, \boldsymbol{\gamma}) &= -\frac{TN}{2}\ln 2\pi - \frac{N\left(T-1\right)}{2}\ln\sigma_e^2 + \frac{N}{2}\ln|\mathbf{C}|^2 - \frac{N}{2}\ln\sigma_\alpha^2 \qquad (5) \\
&\quad -\frac{1}{2}\boldsymbol{\varepsilon}'\left(\mathbf{I}_N \otimes \mathbf{A}^*\right)\boldsymbol{\varepsilon} - \frac{1}{2}\ln\left|\mathbf{I}_T + N\sigma_u^2\mathbf{\Psi}_\lambda\mathbf{A}^*\right| \\
&\quad +\frac{\sigma_u^2}{2}\boldsymbol{\varepsilon}'(\boldsymbol{\iota}_N \otimes \mathbf{A}^*)\left[\mathbf{I}_T + N\sigma_u^2\mathbf{\Psi}_\lambda\mathbf{A}^*\right]^{-1}\mathbf{\Psi}_\lambda(\boldsymbol{\iota}_N' \otimes \mathbf{A}^*)\boldsymbol{\varepsilon}
\end{aligned}$$

where $\boldsymbol{\delta} = \left(\alpha, \boldsymbol{\beta}', \boldsymbol{\pi}', \boldsymbol{\tau}'\right)'$ and $\boldsymbol{\gamma}$ is the vector of covariance parameters, $(\sigma_\mu^2, \sigma_e^2, \rho_v, \sigma_u^2, \rho_\lambda)$.

Evaluation of the likelihood requires numerical computation of the determinant and inverse of the $T \times T$ matrix $\mathbf{I}_T + N\sigma_u^2\mathbf{\Psi}_\lambda\mathbf{A}^*$. The elements of the score for the comprehensive log likelihood (5) are given in appendix A.1 and the information matrix in appendix A.2.

3 Asymptotic properties

Establishing consistency and asymptotic normality is complicated due to the fact that the likelihood contains terms of different orders. Furthermore the likelihood cannot be evaluated analytically which complicates matter further.

3.1 Assumptions

The following assumptions are sufficient for the results

(a) $\mu_i \sim N\left(0, \sigma_\mu^2\right)$, $u_t \sim N\left(0, \sigma_u^2\right)$, $e_{it} \sim N\left(0, \sigma_e^2\right)$ *independent of each other and* $\mathbf{X}$, $\mathbf{d}$ *and* $\mathbf{h}$. *In addition* $\mathbf{X}$, $\mathbf{d}$ *and* $\mathbf{h}$ *have full column ranks* k_1, k_2 *and* k_3 *respectively where* $(\mathbf{X}_{i1}, \ldots, \mathbf{X}_{iT}, \mathbf{h}_i)$ *is iid across* i, $i = 1, \ldots, N$ *and* $(\mathbf{X}_{1t}, \ldots, \mathbf{X}_{Nt}, \mathbf{d}_t)$ *is strictly stationary and ergodic across* t, $t = 1, \ldots, T$ *with* $E\left|X_{jit}\right|^2 < \infty$, $j = 1, \ldots, k_1$, $E\left|d_{lt}\right|^2 < \infty$, $l = 1, \ldots, k_2$ *and* $E\left|h_{si}\right|^2 < \infty$, $s = 1, \ldots, k_3$

(b) $\Theta \equiv \{\theta : \boldsymbol{\delta}'\boldsymbol{\delta} \leq c < \infty, 0 < \sigma_{j,lb}^2 \leq \sigma_j^2 \leq \sigma_{j,ub}^2, -1 < \rho_{i,lb} \leq \rho_i \leq \rho_{i,ub} < 1\}$, *where* ub, lb *denote upper and lower bound respectively and* $j = \mu, u, e$, $i = \lambda, v$ *with* θ_0 *the true parameter vector belonging to the interior of* Θ

(c) *The normalized moment matrix,* $\frac{1}{NT}\mathbf{Z}'\mathbf{Z}$, *converge in probability to a finite positive-definite matrix as* $N \to \infty, T \to \infty$ *or* $N, T \to \infty$. *In addition there exists a diagonal matrix, say* $\boldsymbol{\Upsilon}$, *such that the normalized quadratic form*

$$\boldsymbol{\Upsilon}^{-1}\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z}\boldsymbol{\Upsilon}^{-1}$$

converge in probability uniformly on Θ *to a finite positive-definite matrix as* $N \to \infty, T \to \infty$ *or* $N, T \to \infty$

The normality assumption on μ_i, u_t and e_{it} in (a) is certainly not necessary for consistency arguments. It is well-known that maximizing a normal log-likelihood even though the errors are non-normal will in general give consistent estimates given some moment conditions on μ_i, u_t, e_{it}. Inference is however more complicated so it is convenient to stay in the Gaussian framework.

Assumptions (b) is standard whereas assumption (c) may require some clarification. The first part of assumption (c) is the usual moment condition on the explanatory variables encountered in the asymptotic analysis of least squares models. The second part is concerned with the quadratic form, $\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z}$. It amounts to assuming that the normalized quadratic form, $\boldsymbol{\Upsilon}^{-1}\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z}\boldsymbol{\Upsilon}^{-1}$, have the required limit properties. Lemma 6 in appendix B derives the scalings necessary for the block diagonal elements to converge to positive definite matrices. It follows from this that the scaling matrix must be given by

$$\operatorname{diag} \boldsymbol{\Upsilon} = \left(\min\left(\sqrt{N}, \sqrt{T}\right), \mathbf{F}_\beta, \mathbf{F}_\pi, \mathbf{F}_\tau\right) \tag{6}$$

where $\mathbf{F}_\beta$ is a vector containing $k_1 \sqrt{NT}$, and $\mathbf{F}_\pi, \mathbf{F}_\tau$ are vectors containing $k_2 \sqrt{T}$ and $k_3 \sqrt{N}$ respectively. Contrary to Amemiya (1971) we do not assume that $\operatorname{plim}_{N,T\to\infty} \frac{1}{NT} \left(\boldsymbol{\iota}_{NT}, \mathbf{X}\right)' \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\iota}_{NT}, \mathbf{X}\right)$ is non-singular. This is not true as can be seen from the form of the scaling matrix in (6). The constant needs different normalization and to complicate matters further the appropriate normalization depends on the relative rate of increase of N and T. This indicates a general problem with time-invariant and/or individual-invariant explanatory variables and in this sense we can interpret assumption (c) as that the $\mathbf{H}$ and $\mathbf{D}$ matrices contain variables with "sufficient variation" in the N and T dimension respectively. In fact, $\operatorname{plim}_{T\to\infty} \frac{1}{T}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H}$ and $\operatorname{plim}_{N\to\infty} \frac{1}{N}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D}$ are null matrices whereas $\operatorname{plim}_{T\to\infty} \mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H}$ and $\operatorname{plim}_{N\to\infty} \mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D}$ are random matrices. The appropriate normalizations of these information elements as both N and T grows large are $\frac{1}{N}$ and $\frac{1}{T}$ respectively and in contrast to the constant term these normalizations do not depend on the relative rate of increase of N and T. This illustrates that the behavior of the quadratic form, $\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z}$, may differ sharply from that of the "ordinary form", $\mathbf{Z}'\mathbf{Z}$.

We might remark here that the normalization matrix given in (6) and of course assumption (b) as well is only appropriate for the two-way model. For the one-way model with individual effects $\mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D} = N\mathbf{d}'\mathbf{A}^*\mathbf{d}$ and hence $\operatorname{plim}_{N\to\infty} \frac{1}{N}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D}$ is a random matrix. Similarly in the one-way model with time effects $\operatorname{plim}_{T\to\infty} \frac{1}{T}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H}$ is a random matrix. The appropriate normalizations of the information elements $\mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D}$, $\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H}$ are $\frac{1}{NT}$ and $\frac{1}{N}$ respectively in the one-way model with individual effects and $\frac{1}{T}$ and $\frac{1}{NT}$ respectively in the one-way model with time effects. The unique scaling matrix for the one-way model with individual effects is obtained by letting the first element of $\operatorname{diag} \boldsymbol{\Upsilon}$ be replaced with $\sqrt{N}$ and $\mathbf{F}_\pi$ a vector containing $k_2 \sqrt{NT}$. For the time effects case this matrix is obtained by replacing the first element of $\operatorname{diag} \boldsymbol{\Upsilon}$ with $\sqrt{T}$ and letting $\mathbf{F}_\tau$ be a vector containing $k_3 \sqrt{NT}$.

For the purpose of giving results for the one-way models we define $\Theta^{(i)}$ as the compact parameter space for the parameters of the individual effects model, $\theta^{(i)} = \left(\boldsymbol{\delta}, \boldsymbol{\gamma}^{(i)}\right), \gamma^{(i)} = \left(\sigma_\mu^2, \sigma_e^2, \rho_v\right)$. Correspondingly we define $\Theta^{(t)}$ as the compact parameter space for the parameters of the time effects model, $\theta^{(t)} = \left(\boldsymbol{\delta}, \boldsymbol{\gamma}^{(t)}\right), \boldsymbol{\gamma}^{(t)} = \left(\sigma_e^2, \rho_v, \sigma_u^2, \rho_\lambda\right)$ and make the following additional assumptions

(b$_{(i)}$) $\Theta^{(i)} \equiv \{\theta^{(i)} : \boldsymbol{\delta}'\boldsymbol{\delta} \leq c < \infty, 0 < \sigma_{j,lb}^2 \leq \sigma_j^2 \leq \sigma_{j,ub}^2, -1 < \rho_{v,lb} \leq \rho_v \leq \rho_{v,ub} < 1\}$, *where* ub, lb *denote upper and lower bound respectively and* $j = \mu, e$ *with* $\theta_0^{(i)}$ *the true parameter vector belonging to the interior of* $\Theta^{(i)}$

($b_{(t)}$) $\Theta^{(t)} \equiv \{\theta : \boldsymbol{\delta}'\boldsymbol{\delta} \leq c < \infty, 0 < \sigma^2_{j,lb} \leq \sigma^2_j \leq \sigma^2_{j,ub}, -1 < \rho_{i,lb} \leq \rho_i \leq \rho_{i,ub} < 1\}$, *where* ub, lb *denote upper and lower bound respectively and* $j = u, e$, $i = \lambda, v$ *with* $\theta_0^{(t)}$ *the true parameter vector belonging to the interior of* $\Theta^{(t)}$

Unless otherwise indicated in the following results for the comprehensive model use assumptions (a)-(c) and results for the one-way model with individual effects use assumptions (a), ($b_{(i)}$), (c). Accordingly, results for the one-way model with time effects use assumptions (a), ($b_{(t)}$) and (c).

3.2 Consistency

Our first result is for the comprehensive model specified by the log-likelihood (5). Define $\theta = (\boldsymbol{\delta}, \boldsymbol{\gamma})$ and let $\widehat{\theta} = \left(\widehat{\boldsymbol{\delta}}, \widehat{\boldsymbol{\gamma}}\right)$, $\theta_0 = (\boldsymbol{\delta}_0, \boldsymbol{\gamma}_0)$ denote the estimator and true parameters respectively

Theorem 1 *(Comprehensive model)*

(i) $\widehat{\theta} \xrightarrow{p} \theta_0$ *on* Θ *as* $N, T \to \infty$ *(it does not matter how)*

(ii) $\widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0, \widehat{\boldsymbol{\tau}} \xrightarrow{p} \boldsymbol{\tau}_0$ *on* Θ *as* $N \to \infty$ *and if in addition* $T \geq 2$, $\left(\widehat{\sigma}^2_\mu, \widehat{\sigma}^2_e, \widehat{\rho}_v\right) \xrightarrow{p} \left(\sigma^2_{\mu 0}, \sigma^2_{e0}, \rho_{v0}\right)$ *on* Θ *as* $N \to \infty$

(iii) $\widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$, $\widehat{\boldsymbol{\pi}} \xrightarrow{p} \boldsymbol{\pi}_0$ *on* Θ *as* $T \to \infty$ *and if in addition* $N \geq 2$, $\left(\widehat{\sigma}^2_e, \widehat{\rho}_v, \widehat{\sigma}^2_u, \widehat{\rho}_\lambda\right) \xrightarrow{p} \left(\sigma^2_{e0}, \rho_{v0}, \sigma^2_{u0}, \rho_{\lambda 0}\right)$ *in an open neighborhood of* $\left(\sigma^2_{e0}, \rho_{v0}, \sigma^2_{u0}, \rho_{\lambda 0}\right)$

The proof proceeds by examining the probability limit of the log-likelihood standardized by $\frac{1}{NT}$. This method is not useful for dealing with the constant term but it allows us to prove some global consistency results for the variance parameters which are not easily obtained otherwise. The asymptotic properties of the constant term are essentially established in two lemmas, lemma 6 and lemma 7 given in appendix B. Lemma 6 shows that $\widehat{\alpha} = \min\left(\sqrt{N}, \sqrt{T}\right)$ consistent and hence the constant is not consistently estimated if only N or $T \to \infty$. Note that the inconsistency of the constant does not affect consistency of the $\sqrt{N}$ consistent parameters as $N \to \infty$. Nor does it affect consistency of the $\sqrt{T}$ consistent parameters as $T \to \infty$. The intuition for this is that these estimators do not (at least not asymptotically) use information about the constant. Analogously, inconsistency of for example $\boldsymbol{\pi}$ (the parameters of individual-invariant explanatory variables) as $N \to \infty$ does not affect consistency of the $\sqrt{N}$ consistent parameters[1].

Note that we assumed $T \geq 2$ as $N \to \infty$ to achieve identification of the variance parameters $\left(\sigma_\mu^2, \sigma_e^2, \rho_v\right)$ and $N \geq 2$ as $T \to \infty$ to achieve identification of the variance parameters $(\sigma_e^2, \rho_v, \sigma_u^2, \rho_\lambda)$. A similar requirement appears in assumption (a) and these conditions are frequently redundant when there are time or individual-invariant variables in the model.

A number of special cases emerges from the above theorem. For example, consistency results for the two-way model with serially correlated time effects and the two-way model with serially correlated idiosyncratic errors follow as direct corollaries from theorem 1. In addition if $\rho_v = \rho_\lambda = 0$ and we have no time or individual-invariant explanatory variables theorem 1 (i) gives the consistency result of Amemiya (1971) for the standard two-way model. Theorem 1 (ii) and (iii) then gives consistency results as $N \to \infty$ and $T \to \infty$ respectively not covered in Amemiya (1971)[2].

Theorem 1 does not apply to the one-way model with both serially correlated time effects and serially correlated idiosyncratic errors since we have not allowed for $\sigma_\mu^2 = 0$. Consistency results for this model are however straightforward to obtain

Corollary 1 *(One-way model with time effects)*

(i) $\widehat{\theta}^{(t)} \xrightarrow{p} \theta_0^{(t)}$ *on* $\Theta^{(t)}$ *as* $N, T \to \infty$ *(it does not matter how)*

(ii) $\widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$, $\widehat{\boldsymbol{\tau}} \xrightarrow{p} \boldsymbol{\tau}_0$ *and* $\left(\widehat{\sigma}_e^2, \widehat{\rho}_v\right) \xrightarrow{p} (\sigma_{e0}^2, \rho_{v0})$ *on* $\Theta^{(t)}$ *as* $N \to \infty$

(iii) $\widehat{\alpha} \xrightarrow{p} \alpha_{0,}\widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$, $\widehat{\boldsymbol{\pi}} \xrightarrow{p} \boldsymbol{\pi}_0$ *and* $\widehat{\boldsymbol{\tau}} \xrightarrow{p} \boldsymbol{\tau}_0$ *on* $\Theta^{(t)}$ *as* $T \to \infty$ *and if in addition* $N \geq 2$, $\left(\widehat{\sigma}_e^2, \widehat{\rho}_v, \widehat{\sigma}_u^2, \widehat{\rho}_\lambda\right) \xrightarrow{p} \left(\sigma_{e0}^2, \rho_{v0,}\sigma_{u0}^2, \rho_{\lambda 0}\right)$ *in an open neighborhood of* $\left(\sigma_{e0}^2, \rho_{v0,}\sigma_{u0}^2, \rho_{\lambda 0}\right)$

In contrast to the comprehensive model considered in theorem 1 it is in this case possible to estimate all the parameters consistently as only $T \to \infty$. This follows since there is no individual effect which confounds with the constant term or the time-invariant explanatory variables. The constant is accordingly $\sqrt{T}$ consistent no matter what the relative rate of increase of N and T and $\boldsymbol{\tau}$ is accordingly $\sqrt{NT}$ consistent. The non-presence of individual

[1] The phrase "inconsistent parameters" is used here to refer to parameters whose estimator converge to non-degenerate random variables.

[2] For the standard two-way model it is straightforward to prove global consistency of $\widehat{\sigma}_u^2, \widehat{\sigma}_e^2$ as $T \to \infty$ (assuming $N \geq 2$).

Table 1 Consistency properties of random effects models

Model	Case		Case	
2-way $(\mu_i, \lambda_t, v_{it})$	$N \to \infty$		$T \to \infty$	
	C	NC	C	NC
λ_t, v_{it} iid	$\boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_\mu^2, \sigma_e^2$	$\alpha, \boldsymbol{\pi}, \sigma_u^2$	$\boldsymbol{\beta}, \boldsymbol{\pi}, \sigma_u^2, \sigma_e^2$	$\alpha, \boldsymbol{\tau}, \sigma_\mu^2$
λ_t $AR(1), v_{it}$ iid	$\boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_\mu^2, \sigma_e^2$	$\alpha, \boldsymbol{\pi}, \sigma_u^2, \rho_\lambda$	$\boldsymbol{\beta}, \boldsymbol{\pi}, \sigma_u^2, \rho_\lambda, \sigma_e^2$	$\alpha, \boldsymbol{\tau}, \sigma_\mu^2$
λ_t iid, v_{it} $AR(1)$	$\boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_\mu^2, \sigma_e^2, \rho_v$	$\alpha, \boldsymbol{\pi}, \sigma_u^2$	$\boldsymbol{\beta}, \boldsymbol{\pi}, \sigma_u^2, \sigma_e^2, \rho_v$	$\alpha, \boldsymbol{\tau}, \sigma_\mu^2$
λ_t $AR(1), v_{it}$ $AR(1)$	$\boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_\mu^2, \sigma_e^2, \rho_v$	$\alpha, \boldsymbol{\pi}, \sigma_u^2, \rho_\lambda$	$\boldsymbol{\beta}, \boldsymbol{\pi}, \sigma_u^2, \rho_\lambda, \sigma_e^2, \rho_v$	$\alpha, \boldsymbol{\tau}, \sigma_\mu^2$
1-way (λ_t, v_{it})				
	C	NC	C	NC
λ_t, v_{it} iid	$\boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_e^2$	$\alpha, \boldsymbol{\pi}, \sigma_u^2$	$\alpha, \boldsymbol{\beta}, \boldsymbol{\pi}, \boldsymbol{\tau}, \sigma_u^2, \sigma_e^2$	
λ_t $AR(1), v_{it}$ iid	$\boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_e^2$	$\alpha, \boldsymbol{\pi}, \sigma_u^2, \rho_\lambda$	$\alpha, \boldsymbol{\beta}, \boldsymbol{\pi}, \boldsymbol{\tau}, \sigma_u^2, \rho_\lambda, \sigma_e^2$	
λ_t iid, v_{it} $AR(1)$	$\boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_e^2, \rho_v$	$\alpha, \boldsymbol{\pi}, \sigma_u^2$	$\alpha, \boldsymbol{\beta}, \boldsymbol{\pi}, \boldsymbol{\tau}, \sigma_u^2, \sigma_e^2, \rho_v$	
λ_t $AR(1), v_{it}$ $AR(1)$	$\boldsymbol{\beta}, \boldsymbol{\tau}, \sigma_e^2, \rho_v$	$\alpha, \boldsymbol{\pi}, \sigma_u^2, \rho_\lambda$	$\alpha, \boldsymbol{\beta}, \boldsymbol{\pi}, \boldsymbol{\tau}, \sigma_u^2, \rho_\lambda, \sigma_e^2, \rho_v$	
1-way (μ_i, v_{it})				
	C	NC	C	NC
v_{it} iid	$\alpha, \boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\pi}, \sigma_\mu^2, \sigma_e^2$		$\boldsymbol{\beta}, \boldsymbol{\pi}, \sigma_e^2$	$\alpha, \boldsymbol{\tau}, \sigma_\mu^2$
v_{it} $AR(1)$	$\alpha, \boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\pi}, \sigma_\mu^2, \sigma_e^2, \rho_v$		$\boldsymbol{\beta}, \boldsymbol{\pi}, \sigma_e^2, \rho_v$	$\alpha, \boldsymbol{\tau}, \sigma_\mu^2$

Abbreviations: C=Consistent; NC=Not Consistent

effects also implies that there is a somewhat weaker identification condition on the variance parameters (σ_e^2, ρ_v) as $N \to \infty$.

Finally we give consistency results for the one-way model with individual effects and serially correlated idiosyncratic errors

Corollary 2 *(One-way model with individual effects)*

(i) $\widehat{\theta}^{(i)} \to \theta_0^{(i)}$ *on* $\Theta^{(i)}$ *as* $N, T \to \infty$ *(it does not matter how)*

(ii) $\widehat{\alpha} \xrightarrow{p} \alpha_0, \widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$, $\widehat{\boldsymbol{\tau}} \xrightarrow{p} \boldsymbol{\tau}_0$ *and* $\widehat{\boldsymbol{\pi}} \xrightarrow{p} \boldsymbol{\pi}_0$ *on* $\Theta^{(i)}$ *as* $N \to \infty$ *and if in addition* $T \geq 2$, $\left(\widehat{\sigma}_\mu^2, \widehat{\sigma}_e^2, \widehat{\rho}_v\right) \xrightarrow{p} \left(\sigma_{\mu 0}^2, \sigma_{e0}^2, \rho_{v0}\right)$ *on* $\Theta^{(i)}$ *as* $N \to \infty$

(iii) $\widehat{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$, $\widehat{\boldsymbol{\pi}} \xrightarrow{p} \boldsymbol{\pi}_0$ *and* $\left(\widehat{\sigma}_e^2, \widehat{\rho}_v\right) \xrightarrow{p} \left(\sigma_{e0}^2, \rho_{v0}\right)$ *on* $\Theta^{(i)}$ *as* $T \to \infty$

Since no time effect confounds with the constant or the individual-invariant explanatory variables α and $\boldsymbol{\pi}$ are $\sqrt{N}$ and $\sqrt{NT}$ consistent respectively implying that all parameters are consistently estimated as only $N \to \infty$. We also note that we do not need $N \geq 2$ as $T \to \infty$ to identify the variance parameters (σ_e^2, ρ_v).

The results in theorem 1 and corollaries 1 and 2 covers a number of interesting models commonly used in practice and it is useful to summarize the consistency properties obtained. This is done in Table 1.

3.3 Asymptotic normality

3.3.1 Comprehensive model

In this section our interest centers on the asymptotic distribution of the appropriately scaled maximum likelihood estimator $\widehat{\theta} = \left(\widehat{\boldsymbol{\delta}}, \widehat{\boldsymbol{\gamma}}\right)$. Before the statement of the main theorem it is useful to collect some preliminary results which appear in lemma 6 and 7 in appendix B.

Recall that assumption (c) ensures that the part of the limiting information matrix which belongs to the explanatory variables is a positive-definite matrix as either or both of the indices grow large. In case both N and $T \to \infty$ this limiting matrix, denoted $\mathbf{R}$, is obviously non-stochastic. A moments consideration also reveals that this matrix depends on the behavior of the ratio $\frac{N}{T}$

Lemma 1 *If* $\frac{N}{T} \to \infty$

$$\mathbf{R} = \begin{bmatrix} \frac{(1-\rho_\lambda)^2}{\sigma_u^2} & \mathbf{0} & \frac{(1-\rho_\lambda)^2}{\sigma_u^2} E\mathbf{d}_t' & \mathbf{0} \\ & \mathbf{R}_\mathbf{X} & \mathbf{0} & \mathbf{0} \\ & & \frac{1}{\sigma_u^2} E\mathbf{d}_t^\lambda \mathbf{d}_t^{\lambda\prime} & \mathbf{0} \\ & & & \frac{1}{\sigma_\mu^2}\left(E\mathbf{h}_i\mathbf{h}_i' - E\mathbf{h}_i E\mathbf{h}_i'\right) \end{bmatrix}$$

where $\mathbf{R}_\mathbf{X} = \mathrm{plim}_{N,T\to\infty} \frac{1}{NT}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}$ *and* $\mathbf{d}_t^\lambda = (\mathbf{d}_t - \rho_\lambda \mathbf{d}_{t-1})$.
If $\frac{T}{N} \to \infty$

$$\mathbf{R} = \begin{bmatrix} \frac{1}{\sigma_\mu^2} & \mathbf{0} & \mathbf{0} & \frac{1}{\sigma_\mu^2} E\mathbf{h}_i' \\ & \mathbf{R}_\mathbf{X} & \mathbf{0} & \mathbf{0} \\ & & \frac{1}{\sigma_u^2} E\mathbf{d}_t^\lambda \mathbf{d}_t^{\lambda\prime} - \frac{(1-\rho_\lambda)^2}{\sigma_u^2} E\mathbf{d}_t E\mathbf{d}_t' & \mathbf{0} \\ & & & \frac{1}{\sigma_\mu^2} E\mathbf{h}_i\mathbf{h}_i' \end{bmatrix}$$

Finally, if $N, T \to \infty$ *simultaneously*

$$\mathbf{R} = \begin{bmatrix} \omega & \mathbf{0} & \omega E\mathbf{d}_t' & \omega E\mathbf{h}_i' \\ & \mathbf{R}_\mathbf{X} & \mathbf{0} & \mathbf{0} \\ & & \frac{1}{\sigma_u^2} E\mathbf{d}_t^\lambda \mathbf{d}_t^{\lambda\prime} + \upsilon_1 E\mathbf{d}_t E\mathbf{d}_t' & \omega E\mathbf{d}_t E\mathbf{h}_i' \\ & & & \frac{1}{\sigma_\mu^2} E\mathbf{h}_i\mathbf{h}_i' + \upsilon_2 E\mathbf{h}_i E\mathbf{h}_i' \end{bmatrix}$$

where $\omega = \frac{1}{\sigma_\mu^2 + (1-\rho_\lambda)^{-2}\sigma_u^2}$, $\upsilon_1 = \omega - \frac{(1-\rho_\lambda)^2}{\sigma_u^2}$ *and* $\upsilon_2 = \omega - \sigma_\mu^{-2}$.

The lemma shows that when both N and $T \to \infty$ the variance formula, and hence the amount of information in the sample, depends on the behavior of the ratio $\frac{N}{T}$. It is important to notice that this result does not relate to assumptions about the sampling behavior of the time-invariant and the individual-invariant explanatory variables. In fact it holds regardless of wether the time-invariant and the individual-invariant explanatory variables are regarded as fixed or stochastic. Note however that if these variables are centered then $\mathbf{R}$ reduces to a block-diagonal matrix that only depends on the behavior of the ratio $\frac{N}{T}$ through the constant term.

If only N or $T \to \infty$ as in theorem 1 (ii) and (iii) part of the parameter vector is not consistently estimated. In fact, only the subvectors $\theta^i = \left(\boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\gamma}^{(i)}\right), \boldsymbol{\gamma}^{(i)} = (\sigma_\mu^2, \sigma_e^2, \rho_v)$ and $\theta^t = \left(\boldsymbol{\beta}, \boldsymbol{\pi}, \boldsymbol{\gamma}^{(t)}\right)$, $\boldsymbol{\gamma}^{(t)} = (\sigma_e^2, \rho_v, \sigma_u^2, \rho_\lambda)$ are consistently estimated as N and $T \to \infty$ respectively. The limiting distributions of the subsets of consistently estimated parameters are of course only interesting if they are information block-diagonal to the inconsistent parameters. The following lemma, which is a direct consequence of lemma 7 in appendix B, is useful in this respect

Lemma 2 *As $N \to \infty$ the information matrix is block-diagonal between θ^i and $(\alpha, \boldsymbol{\pi}', \sigma_u^2, \rho_\lambda)$ and as $T \to \infty$ the information matrix is block-diagonal between θ^t and $\left(\alpha, \boldsymbol{\tau}', \sigma_\mu^2\right)$*

Motivated by this lemma the theorem below applies a mean-value expansion to the part of the score vector which belongs to the consistent subvectors. In addition the elements of the limiting information matrix relating to the consistently estimated subvectors does not depend on the nuisance parameters $(\alpha, \boldsymbol{\pi}', \sigma_u^2, \rho_\lambda)$ as $N \to \infty$ nor on the nuisance parameters $\left(\alpha, \boldsymbol{\tau}', \sigma_\mu^2\right)$ as $T \to \infty$. This fact is important since it implies that we can obtain useful approximate variance formulas for the subsets of consistently estimated parameters.

We now obtain the main result of this section. For this purpose define $\mathbf{F}_{NT}$, $\mathbf{F}_N$ and $\mathbf{F}_T$ as diagonal matrices with

$$\begin{aligned}
\operatorname{diag} \mathbf{F}_{NT} &= \left\{\min\left(\sqrt{N}, \sqrt{T}\right), \mathbf{F}_\beta, \mathbf{F}_\pi, \mathbf{F}_\tau, \sqrt{N}, \sqrt{NT}, \sqrt{NT}, \sqrt{T}, \sqrt{T}\right\} \\
\operatorname{diag} \mathbf{F}_N &= \left\{\mathbf{F}_\beta, \mathbf{F}_\tau, \sqrt{N}, \sqrt{NT}, \sqrt{NT}\right\} \\
\operatorname{diag} \mathbf{F}_T &= \left\{\mathbf{F}_\beta, \mathbf{F}_\pi, \sqrt{NT}, \sqrt{NT}, \sqrt{T}, \sqrt{T}\right\}
\end{aligned}$$

We shall also need notation for limits of submatrices of the quadratic form in assumption (c). Let $\mathbf{Z}_N = (\mathbf{X}, \mathbf{H})$, $\mathbf{Z}_T = (\mathbf{X}, \mathbf{D})$, $\boldsymbol{\Upsilon}_N$ and $\boldsymbol{\Upsilon}_T$ diagonal such that

$$\begin{aligned}
\operatorname{diag} \boldsymbol{\Upsilon}_N &= (\mathbf{F}_\beta, \mathbf{F}_\tau) \\
\operatorname{diag} \boldsymbol{\Upsilon}_T &= (\mathbf{F}_\beta, \mathbf{F}_\pi)
\end{aligned}$$

with

$$\begin{aligned}
\underset{N\to\infty}{\text{plim}}\, \boldsymbol{\Upsilon}_N^{-1}\mathbf{Z}_N'\boldsymbol{\Sigma}^{-1}\mathbf{Z}_N\boldsymbol{\Upsilon}_N^{-1} &= \mathbf{R}_N \\
\underset{T\to\infty}{\text{plim}}\, \boldsymbol{\Upsilon}_T^{-1}\mathbf{Z}_T'\boldsymbol{\Sigma}^{-1}\mathbf{Z}_T\boldsymbol{\Upsilon}_T^{-1} &= \mathbf{R}_T
\end{aligned}$$

and we further let $\overline{\theta}$ denote a sequence such that $\text{plim}\,\overline{\theta} = \theta_0$

Theorem 2 *(Comprehensive model)*

(i) $\mathbf{F}_{NT}\left(\widehat{\theta}-\theta_0\right) \xrightarrow{d} N\left(\mathbf{0},\mathbf{V}\left(\theta_0\right)\right)$ *as* $N,T\to\infty$*, where*

$$\mathbf{V}^{-1}\left(\theta_0\right) = -\underset{N,T\to\infty}{\text{plim}}\left[\mathbf{F}_{NT}^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta},\boldsymbol{\gamma})}{\partial\theta\partial\theta'}|_{\overline{\theta}}\right)\mathbf{F}_{NT}^{-1}\right] = \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}^{-1}\left(\theta_0\right)_{\gamma} \end{bmatrix}$$

a finite non-singular matrix, with $\mathbf{R}=\mathbf{R}\left(\theta_0\right)$ *a* $\sum_{i=1}^{3} k_i+1$ *dimensional matrix given in lemma 1 and* $\mathbf{V}^{-1}\left(\theta_0\right)_{\gamma}$ *is a diagonal matrix with*

$$\text{diag}\,\mathbf{V}^{-1}\left(\theta_0\right)_{\gamma} = \left\{\frac{1}{2\sigma_{\mu0}^4},\frac{1}{2\sigma_{e0}^4},\frac{1}{\left(1-\rho_{v0}^2\right)},\frac{1}{2\sigma_{u0}^4},\frac{1}{\left(1-\rho_{\lambda0}^2\right)}\right\}$$

(ii) $\mathbf{F}_N\left(\widehat{\theta}^i-\theta_0^i\right) \xrightarrow{d} N\left(\mathbf{0},\mathbf{V}_N\left(\theta_0^i\right)\right)$ *as* $N\to\infty$ *(assuming* $T\geq 2$*), where*

$$\mathbf{V}_N^{-1}\left(\theta_0^i\right) = -\underset{N\to\infty}{\text{plim}}\left[\mathbf{F}_N^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta},\boldsymbol{\gamma})}{\partial\theta^i\partial\theta^{i\prime}}|_{\overline{\theta}^i}\right)\mathbf{F}_N^{-1}\right] = \begin{bmatrix} \mathbf{R}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_N^{-1}\left(\theta_0^i\right)_{\gamma^{(i)}} \end{bmatrix}$$

a finite non-singular matrix, with $\mathbf{R}_N=\mathbf{R}_N\left(\theta_0^i\right)$ *a* k_1+k_3 *dimensional matrix and* $\mathbf{V}_N^{-1}\left(\theta_0^i\right)_{\gamma^{(i)}}$ *given by*

$$\begin{aligned}
&\mathbf{V}_N^{-1}\left(\theta_0^i\right)_{\gamma^{(i)}} \\
&= \frac{1}{2}\begin{bmatrix} \left(\frac{\sigma_{\varpi}^2}{\sigma_{\alpha0}^2\sigma_{\mu0}^2}\right)^2 & \frac{(1-\rho_{v0})^2}{\sigma_{\alpha0}^2\sqrt{T}}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\boldsymbol{\Psi}_v\boldsymbol{\iota}_T^{\alpha} & \frac{\sigma_{e0}^2(1-\rho_{v0})^2}{\sigma_{\alpha0}^2\sqrt{T}}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{L}_v\boldsymbol{\iota}_T^{\alpha} \\
 & \frac{1}{T}\left(\sigma_{\alpha0}^{-4}+(T-1)\,\sigma_{e0}^{-4}\right) & \frac{\sigma_{e0}^2}{T}\,\text{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{L}_v\right) \\
 & & \frac{\sigma_{e0}^4}{T}\,\text{tr}\left(\mathbf{A}^*\mathbf{L}_v\right)^2
\end{bmatrix}
\end{aligned}$$

where $\sigma_{\varpi}^2=\left(\sigma_{\alpha0}^2-\sigma_{e0}^2\right)$

(iii) $\mathbf{F}_T\left(\widehat{\theta}^t-\theta_0^t\right) \xrightarrow{d} N\left(\mathbf{0},\mathbf{V}_T\left(\theta_0^t\right)\right)$ *as* $T\to\infty$ *(assuming* $N\geq 2$*), where*

$$\mathbf{V}_T^{-1}\left(\theta_0^t\right) = -\underset{T\to\infty}{\text{plim}}\left[\mathbf{F}_T^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta},\boldsymbol{\gamma})}{\partial\theta^t\partial\theta^{t\prime}}|_{\overline{\theta}^t}\right)\mathbf{F}_T^{-1}\right] = \begin{bmatrix} \mathbf{R}_T & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_T^{-1}\left(\theta_0^t\right)_{\gamma^{(t)}} \end{bmatrix}$$

a finite non-singular matrix, with $\mathbf{R}_T = \mathbf{R}_T\left(\theta_0^t\right)$ *a* $k_1 + k_2$ *dimensional matrix and* $\mathbf{V}_T^{-1}\left(\theta_0^t\right)_{\gamma^{(t)}}$ *given by*

$$
\begin{aligned}
& \mathbf{V}_T^{-1}\left(\theta_0^t\right)_{\gamma^{(t)}} \\
= \quad & \lim_{T\to\infty} \frac{1}{2T} \begin{bmatrix} \frac{1}{\sigma_{e0}^4}\mathbf{V}^+_{\mathbf{\Psi}_v,\mathbf{\Psi}_v} & \frac{1}{\sigma_{e0}^2}\mathbf{V}^+_{\mathbf{L}_v,\mathbf{\Psi}_v} & \frac{\sqrt{N}}{\sigma_{e0}^4}\mathbf{V}_{\mathbf{\Psi}_v,\mathbf{\Psi}_\lambda} & \frac{\sigma_u^2\sqrt{N}}{\sigma_{e0}^4}\mathbf{V}_{\mathbf{\Psi}_v,\mathbf{L}_\lambda} \\ & \mathbf{V}^+_{\mathbf{L}_v,\mathbf{L}_v} & \frac{\sqrt{N}}{\sigma_{e0}^2}\mathbf{V}_{\mathbf{L}_v,\mathbf{\Psi}_\lambda} & \frac{\sigma_{u0}^2\sqrt{N}}{\sigma_{e0}^2}\mathbf{V}_{\mathbf{L}_v,\mathbf{L}_\lambda} \\ & & \frac{N^2}{\sigma_{e0}^4}\mathbf{V}_{\mathbf{\Psi}_\lambda,\mathbf{\Psi}_\lambda} & \frac{\sigma_{u0}^2 N^2}{\sigma_{e0}^4}\mathbf{V}_{\mathbf{L}_\lambda,\mathbf{\Psi}_\lambda} \\ & & & \frac{\sigma_{u0}^4 N^2}{\sigma_{e0}^4}\mathbf{V}_{\mathbf{L}_\lambda,\mathbf{L}_\lambda} \end{bmatrix}
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{V}_{\mathbf{F},\mathbf{P}} &= \operatorname{tr}\left(\left(\mathbf{\Psi}_v^{-1}\mathbf{F}\mathbf{\Psi}_v^{-1}\mathbf{P}\right)\left(\mathbf{I}_T - 2\mathbf{M}\right)\right) + \operatorname{tr}\left(\mathbf{\Psi}_v^{-1}\mathbf{F}\mathbf{M}\mathbf{\Psi}_v^{-1}\mathbf{P}\mathbf{M}\right) \\
\mathbf{V}^+_{\mathbf{F},\mathbf{P}} &= \operatorname{tr}\left(\left(\mathbf{\Psi}_v^{-1}\mathbf{F}\mathbf{\Psi}_v^{-1}\mathbf{P}\right)\left(\mathbf{I}_T - \frac{2}{N}\mathbf{M}\right)\right) + \frac{1}{N}\operatorname{tr}\left(\mathbf{\Psi}_v^{-1}\mathbf{F}\mathbf{M}\mathbf{\Psi}_v^{-1}\mathbf{P}\mathbf{M}\right)
\end{aligned}
$$

and $\mathbf{M} = \left(\mathbf{I}_T + \frac{\sigma_{e0}^2}{N\sigma_{u0}^2}\mathbf{\Psi}_\lambda\mathbf{\Psi}_v\mathbf{\Psi}_\lambda^{-2}\right)^{-1}$

Corresponding asymptotic normality results for the standard two-way model and the two-way model with serially correlated time effects or serially correlated idiosyncratic errors follow directly from theorem 2.

3.3.2 One-way models

Asymptotic normality results for the one-way models considered in corollary 1 and 2 can be derived quite easily given theorem 2. We concentrate on the one-way model with individual effects in this section, corresponding qualitative results for the one-way model with time effects follow similarly.

The limiting information matrix is, as in the two-way model, block-diagonal between consistent and inconsistent parameters. This allows us to obtain the marginal limiting distribution of the consistently estimated parameters when $T \to \infty$ in the same manner as for the two-way model. Also, the limiting information matrix for the consistently estimated parameters does not depend on the inconsistent nuisance parameters, ensuring that we can estimate the limiting variance consistently in the $T \to \infty$ case. In contrast to the two-way model all parameters are consistent as $N \to \infty$ and we obtain joint asymptotic normality for the full parameter vector under $N \to \infty$ as well as $N, T \to \infty$.

Make the following definitions

$$
\begin{aligned}
\operatorname{diag}\mathbf{F}_{NT}^{(i)} &= \left\{\sqrt{N}, \mathbf{F}_\beta, \mathbf{F}_\pi, \mathbf{F}_\tau, \sqrt{N}, \sqrt{NT}, \sqrt{NT}\right\} \\
\operatorname{diag}\mathbf{F}_{T}^{(i)} &= \left\{\mathbf{F}_\beta, \mathbf{F}_\pi, \sqrt{NT}, \sqrt{NT}\right\}
\end{aligned}
$$

where $\mathbf{F}_{NT}^{(i)}, \mathbf{F}_{T}^{(i)}$ are diagonal matrices and $\mathbf{F}_{\pi}$ is as in assumption ($\mathrm{c}_{(i)}$). We also define $\theta^{(i)t} = \left(\boldsymbol{\beta}, \boldsymbol{\pi}, \boldsymbol{\gamma}^{(i)t}\right)$, $\boldsymbol{\gamma}^{(i)t} = (\sigma_e^2, \rho_v)$ and with some further obvious notation we have

Corollary 3 *(One-way model with individual effects)*

(i) $\mathbf{F}_{NT}^{(i)}\left(\widehat{\theta}^{(i)} - \theta_0^{(i)}\right) \xrightarrow{d} N\left(\mathbf{0}, \mathbf{V}^{(i)}\left(\theta_0^{(i)}\right)\right)$ *as* $N, T \to \infty$, *where*

$$\begin{aligned}\left[\mathbf{V}^{(i)}\left(\theta_0^{(i)}\right)\right]^{-1} &= -\operatorname*{plim}_{N,T\to\infty}\left[\left(\mathbf{F}_{NT}^{(i)}\right)^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma}^{(i)})}{\partial\theta^{(i)}\partial\theta^{(i)\prime}}|_{\overline{\theta}^{(i)}}\right)\left(\mathbf{F}_{NT}^{(i)}\right)^{-1}\right] \\ &= \begin{bmatrix} \mathbf{R}^{(i)} & \mathbf{0} \\ \mathbf{0} & \left[\mathbf{V}^{(i)}\left(\theta_0^{(i)}\right)_{\gamma^{(i)}}\right]^{-1}\end{bmatrix}\end{aligned}$$

a finite non-singular matrix, with $\mathbf{R}^{(i)} = \mathbf{R}^{(i)}\left(\theta_0^{(i)}\right)$ *a* $\sum_{i=1}^{3} k_i + 1$ *dimensional matrix given by*

$$\mathbf{R}^{(i)} = \begin{bmatrix} \frac{1}{\sigma_\mu^2} & \mathbf{0} & \mathbf{0} & \frac{1}{\sigma_\mu^2}E\mathbf{h}_i \\ & \mathbf{R}_{\mathbf{X}}^{(i)} & \mathbf{R}_{\mathbf{X},\mathbf{D}}^{(i)} & \mathbf{0} \\ & & \frac{1}{\sigma_e^2}E\mathbf{d}_t^{v\prime}\mathbf{d}_t^{v} - \frac{(1-\rho_v)^2}{\sigma_e^2}E\mathbf{d}_t^{\prime}E\mathbf{d}_t & \mathbf{0} \\ & & & \frac{1}{\sigma_\mu^2}E\mathbf{h}_i^{\prime}\mathbf{h}_i \end{bmatrix}$$

where $\mathbf{R}_{\mathbf{X}}^{(i)} = \operatorname{plim}_{N,T\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i^{\prime}\mathbf{A}^{*}\mathbf{X}_i$, $\mathbf{R}_{\mathbf{X},\mathbf{D}}^{(i)} = \operatorname{plim}_{N,T\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i^{\prime}\mathbf{A}^{*}\mathbf{d}$ *and* $\mathbf{d}_t^{v} = (\mathbf{d}_t - \rho_v\mathbf{d}_{t-1})$ *and* $\left[\mathbf{V}^{(i)}\left(\theta_0^{(i)}\right)_{\gamma^{(i)}}\right]^{-1}$ *is a diagonal matrix with*

$$\operatorname{diag}\left[\mathbf{V}^{(i)}\left(\theta_0^{(i)}\right)_{\gamma^{(i)}}\right]^{-1} = \left\{\frac{1}{2\sigma_{\mu 0}^4}, \frac{1}{2\sigma_{e0}^4}, \frac{1}{(1-\rho_{v0}^2)}\right\}$$

(ii) $\mathbf{F}_{NT}^{(i)}\left(\widehat{\theta}^{(i)} - \theta_0^{(i)}\right) \xrightarrow{d} N\left(\mathbf{0}, \mathbf{V}_N^{(i)}\left(\theta_0^{(i)}\right)\right)$ *as* $N \to \infty$ *(assuming* $T \geq 2$*), where*

$$\begin{aligned}\left[\mathbf{V}_N^{(i)}\left(\theta_0^{(i)}\right)\right]^{-1} &= -\operatorname*{plim}_{N\to\infty}\left[\left(\mathbf{F}_{NT}^{(i)}\right)^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma}^{(i)})}{\partial\theta^{(i)}\partial\theta^{(i)\prime}}|_{\overline{\theta}^{(i)}}\right)\left(\mathbf{F}_{NT}^{(i)}\right)^{-1}\right] \\ &= \begin{bmatrix} \mathbf{R}_N^{(i)} & \mathbf{0} \\ \mathbf{0} & \left[\mathbf{V}_N^{(i)}\left(\theta_0^{(i)}\right)_{\gamma^{(i)}}\right]^{-1}\end{bmatrix}\end{aligned}$$

a finite non-singular matrix, with $\mathbf{R}_N^{(i)} = \mathbf{R}_N^{(i)}\left(\theta_0^{(i)}\right)$ *a* $\sum_{i=1}^{3} k_i + 1$ *dimensional matrix and* $\left[\mathbf{V}_N^{(i)}\left(\theta_0^{(i)}\right)_{\gamma^{(i)}}\right]^{-1}$ *given by*

$$\begin{aligned}
&\left[\mathbf{V}_N^{(i)}\left(\theta_0^{(i)}\right)_{\gamma^{(i)}}\right]^{-1} \\
&= \frac{1}{2}\begin{bmatrix} \left(\frac{\sigma_{\varpi}^2}{\sigma_{\alpha 0}^2\sigma_{\mu 0}^2}\right)^2 & \frac{(1-\rho_{v0})^2}{\sigma_{\alpha 0}^2\sqrt{T}}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\boldsymbol{\Psi}_v\boldsymbol{\iota}_T^{\alpha} & \frac{\sigma_{e0}^2(1-\rho_{v0})^2}{\sigma_{\alpha 0}^2\sqrt{T}}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{L}_v\boldsymbol{\iota}_T^{\alpha} \\ & \frac{1}{T}\left(\sigma_{\alpha 0}^{-4} + (T-1)\,\sigma_{e0}^{-4}\right) & \frac{\sigma_{e0}^2}{T}\operatorname{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{L}_v\right) \\ & & \frac{\sigma_{e0}^4}{T}\operatorname{tr}\left(\mathbf{A}^*\mathbf{L}_v\right)^2 \end{bmatrix}
\end{aligned}$$

where $\sigma_{\varpi}^2 = \left(\sigma_{\alpha 0}^2 - \sigma_{e0}^2\right)$

(iii) $\mathbf{F}_T^{(i)}\left(\widehat{\theta}^{(i)t} - \theta_0^{(i)t}\right) \xrightarrow{d} N\left(\mathbf{0}, \mathbf{V}_T^{(i)}\left(\theta_0^{(i)t}\right)\right)$ *as* $T \to \infty$, *where*

$$\begin{aligned}
\left[\mathbf{V}_T^{(i)}\left(\theta_0^{(i)t}\right)\right]^{-1} &= -\operatorname*{plim}_{T\to\infty}\left[\left(\mathbf{F}_T^{(i)}\right)^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma}^{(i)t})}{\partial\theta^{(i)t}\partial\theta^{(i)t\prime}}|_{\overline{\theta}^{(i)t}}\right)\left(\mathbf{F}_T^{(i)}\right)^{-1}\right] \\
&= \begin{bmatrix} \mathbf{R}_T^{(i)} & \mathbf{0} \\ \mathbf{0} & \left[\mathbf{V}_T^{(i)}\left(\theta_0^{(i)t}\right)_{\gamma^{(i)t}}\right]^{-1} \end{bmatrix}
\end{aligned}$$

a finite non-singular matrix, with $\mathbf{R}_T^{(i)} = \mathbf{R}_T^{(i)}\left(\theta_0^{(i)t}\right)$ *a* $k_1 + k_2$ *dimensional matrix and* $\left[\mathbf{V}_T^{(i)}\left(\theta_0^{(i)t}\right)_{\gamma^{(i)t}}\right]^{-1}$ *is a diagonal matrix with*

$$\operatorname{diag}\left[\mathbf{V}_T^{(i)}\left(\theta_0^{(i)t}\right)_{\gamma^{(i)t}}\right]^{-1} = \left\{\frac{1}{2\sigma_{e0}^4}, \frac{1}{\left(1-\rho_{v0}^2\right)}\right\}$$

Comparing the results in the corollary above to the results in theorem 2 we note that

Property 1 *In contrast to the comprehensive model the information elements of* $\alpha, \boldsymbol{\pi}$ *and* $\boldsymbol{\tau}$ *does not depend on the behavior of the ratio* $\frac{N}{T}$ *as both* N *and* T *grows large*

Property 2 *As* $N \to \infty$ *(or* $N, T \to \infty$*) the variance matrix of the variance parameters* $\gamma^{(i)}$ *is the same in both models*

That is we have the same large N asymptotics for the variance parameters, $\gamma^{(i)}$ in the one-way model with individual effects and the two-way model. Noting that the one-way model with individual specific random effects is typically used in situations where large N asymptotics are appropriate this indicates that it is asymptotically costless to variance robustify by including time specific random effects as well. If in addition $\mathbf{h}$ is centered and $\mathbf{X}$ is centered in the N dimension we have the same large N limiting variance in these models for the parameter vectors $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$ as well.

3.4 Misspecification

It is well-known that in the framework of the classical linear model misspecification of the variance does in general not affect consistency of the regression parameters, only efficiency. Unfortunately, in the present situation this need not be true. As indicated by the results in theorem 1 and corollaries 1 and 2 problems arise since the true and the perceived error component structure need not agree on the appropriate probabilistic orders[3]. The theorem below illustrates what can happen

Theorem 3 (Misspecification of error components) *Suppose assumptions $(a), (b)$ and (c) holds and the true model is the comprehensive model considered in theorem 1 but the estimated model is the one-way model with individual effects considered in corollary 2. Then, for $k_i = 1, i = 1, 2, 3$*

(i) *$\widehat{\gamma}^{(i)t}$ is inconsistent as $N \to \infty, T \to \infty$ or $N, T \to \infty$ and $\widehat{\sigma}^2_\mu$ is inconsistent as $N \to \infty$*

(ii) *As both $N, T \to \infty$ (it does not matter how) $\widehat{\boldsymbol{\delta}} \xrightarrow{p} \boldsymbol{\delta}_0$ on Θ and as $T \to \infty$ $\left(\widehat{\beta}, \widehat{\pi}\right) \xrightarrow{p} (\beta_0, \pi_0)$ on Θ whereas $\widehat{\alpha}$ and $\widehat{\tau}$ are inconsistent. In case of $N \to \infty$ both $(\widehat{\alpha}, \widehat{\pi})$ are inconsistent, the situation for $\left(\widehat{\beta}, \widehat{\tau}\right)$ requires us to distinguish between if $\operatorname{plim}_{N\to\infty} \frac{1}{N} \sum X_{it} h_i$ is zero or not. If non-zero, $\widehat{\beta} \xrightarrow{p} \beta_0$ or $\widehat{\tau} \xrightarrow{p} \tau_0$ (or both) on Θ iff $\mathbf{h}$ is centered and $\mathbf{X}$ is centered in the N dimension. If zero, $\widehat{\beta} \xrightarrow{p} \beta_0$ on Θ iff $\mathbf{X}$ is centered in the N dimension and $\widehat{\tau} \xrightarrow{p} \tau_0$ on Θ iff $\mathbf{h}$ is centered*

Part two of the theorem might seem counterintuitive in the light of standard theory for linear regression. The key to understanding the result is to

[3] Misspecification of the error components imply that the variance of the score and the negative expected hessian need not be equal. In addition they need not have the same probabilistic orders.

note that it is the case $N, T \to \infty$, where all regression parameters are estimated consistently, that corresponds to the standard theory. In the $N \to \infty$ case we may think of the time effects as dummy variables erroneously excluded from the model. Consistent estimation of the remaining regression parameters then requires that the corresponding explanatory variables are orthogonal to the excluded variables, hence the need for centering. Although centering of the data recovers some of the consistency properties for the mean parameters of a correctly specified one-way model it does not, and in contrast to the robustification result in property 2, lead to the same asymptotic distribution[4]. There is a loss of efficiency and a sandwich-type variance-covariance estimator should be used since the information matrix equality fails to hold. Also note that the driving force for the result is the presence of the time specific effects per se. Theorem 3 holds wether λ_t is serially correlated or not.

4 Final remarks

Panel data models which allow for serial correlation are extensively used in applied econometrics. This paper has explored the large sample theory for a comprehensive specification which nests most of the models used in practice[5]. In contrast to the previous literature we have treated the constant term appropriately as well as allowed for both time or individual-invariant random variables.

In terms of the consistency properties obtained our results reveal an interesting and, perhaps, unexpected difference between ordinary explanatory variables and explanatory variables that are time or individual-invariant. Whereas the parameters of ordinary explanatory variables are always estimated consistently whenever N or $T \to \infty$ the consistency properties of the parameters of time or individual-invariant explanatory variables depend crucially on the model. The source of this difference was attributed to confounding with time effects and/or individual effects and, of course, if there are neither individual nor time effects these parameters have the desirable properties of the parameters of ordinary explanatory variables.

[4] The reader may notice that although we are only able to recover some of the consistency properties of the mean parameters in a correctly specified one-way model we obtain exactly the same consistency properties of the mean parameters as for the two-way model.

[5] Of course, none of the results in this paper are special to models with serial correlation. In addition the results for the variance components and the ordinary explanatory variables do not depend on the presence or non-presence of individual or time-invariant explanatory variables.

Our results on asymptotic normality revealed a useful characterization of the limiting information matrix. The set of consistent parameters (as N or $T \rightarrow \infty$) are information block-diagonal to the set of inconsistent parameters and the set of consistent mean parameters are always information block-diagonal to the set of consistent variance parameters. In addition the elements of information of the consistent parameters do not depend on the inconsistent parameters, ensuring that the variance matrix of consistently estimated parameters can be consistently estimated.

As an application of the results obtained we considered the consequences of error component misspecification. In this situation it is useful to work with deviations from means to guard (incompletely) against possible inconsistency of the mean parameters and indeed the idea of centering is also useful in the context of robustification.

Possible extensions of our results include introducing dynamics in form of a lagged dependent variable as well as allowing for time trends commonly employed in practice. Given the present results one would suspect that a linear time trend is $T^{3/2}$ consistent in the two-way model and the one-way model with time effects but $\sqrt{N}T^{3/2}$ consistent in the one-way model with individual effects. However these and other issues are left for future work.

A Score and Information

A.1 The score vector

This appendix derives the elements of the score vector. For the regression parameters we have the standard result

$$\frac{\partial l}{\partial \boldsymbol{\delta}} = \mathbf{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}$$

and for the variance parameters the score is given by

$$\frac{\partial l}{\partial \gamma_i} = -\frac{1}{2}\,\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \gamma_i}) + \frac{1}{2}\boldsymbol{\varepsilon}'\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \gamma_i}\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}$$

where $\boldsymbol{\gamma} =(\sigma_\mu^2, \sigma_e^2, \rho_v, \sigma_u^2, \rho_\lambda)$

For σ_μ^2 we have

$$\begin{aligned}
\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_\mu^2}\right) &= \mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\left(\mathbf{I}_N \otimes \mathbf{J}_T\right)\right) \\
&= \mathrm{tr}\left(\mathbf{I}_N \otimes \mathbf{A}^*\mathbf{J}_T\right) - \mathrm{tr}\left[\left(\boldsymbol{\iota}_N \otimes \mathbf{A}^*\right)\mathbf{B}^{-1}\left(\boldsymbol{\iota}_N' \otimes \mathbf{A}^*\mathbf{J}_T\right)\right] \\
&= \frac{N\left(1-\rho_v\right)^2 d^2}{\sigma_\alpha^2} - \frac{N\left(1-\rho_v\right)^2}{\sigma_\alpha^2}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}
\end{aligned}$$

where $\mathbf{B}^{-1} = \sigma_u^2\left(\mathbf{I}_T + N\sigma_u^2\boldsymbol{\Psi}_\lambda\mathbf{A}^*\right)^{-1}\boldsymbol{\Psi}_\lambda$

$$\begin{aligned}
\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \sigma_\mu^2}\boldsymbol{\Sigma}^{-1} &= \boldsymbol{\Sigma}^{-1}(\mathbf{I}_N \otimes \mathbf{J}_T)\boldsymbol{\Sigma}^{-1} \\
&= (\mathbf{I}_N \otimes \mathbf{A}^*\mathbf{J}_T\mathbf{A}^*) - (\mathbf{J}_N \otimes \mathbf{A}^*\mathbf{J}_T\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*) \\
&\quad -(\mathbf{J}_N \otimes \mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{J}_T\mathbf{A}^*) \\
&\quad +N(\mathbf{J}_N \otimes \mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{J}_T\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*)
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{\partial l}{\partial \sigma_\mu^2} &= -\frac{N\left(1-\rho_v\right)^2 d^2}{2\sigma_\alpha^2} + \frac{N\left(1-\rho_v\right)^2}{2\sigma_\alpha^2}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha} \\
&\quad +\frac{1}{2}\boldsymbol{\varepsilon}'\left(\mathbf{I}_N \otimes \mathbf{A}^*\mathbf{J}_T\mathbf{A}^*\right)\boldsymbol{\varepsilon} - \widetilde{\boldsymbol{\varepsilon}}'\mathbf{J}_T\overline{\boldsymbol{\varepsilon}} + \frac{N}{2}\overline{\boldsymbol{\varepsilon}}'\mathbf{J}_T\overline{\boldsymbol{\varepsilon}}
\end{aligned}$$

where $\widetilde{\boldsymbol{\varepsilon}} = \left(\boldsymbol{\iota}_N' \otimes \mathbf{A}^*\right)\boldsymbol{\varepsilon}$ and $\overline{\boldsymbol{\varepsilon}} = \left(\boldsymbol{\iota}_N' \otimes \mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\right)\boldsymbol{\varepsilon}$. For σ_e^2 we have

$$
\begin{aligned}
\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial \sigma_e^2}\right) &= \operatorname{tr}\left(\mathbf{\Sigma}^{-1}\left(\mathbf{I}_N \otimes \mathbf{\Psi}_v\right)\right) \\
&= \operatorname{tr}\left(\mathbf{I}_N \otimes \mathbf{A}^*\mathbf{\Psi}_v\right) - \operatorname{tr}\left[\left(\boldsymbol{\iota}_N \otimes \mathbf{A}^*\right)\mathbf{B}^{-1}\left(\boldsymbol{\iota}_N' \otimes \mathbf{A}^*\mathbf{\Psi}_v\right)\right] \\
&= N\operatorname{tr}\left(\mathbf{A}^*\mathbf{\Psi}_v\right) - N\operatorname{tr}\left(\mathbf{A}^*\mathbf{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial \sigma_e^2}\mathbf{\Sigma}^{-1} &= \mathbf{\Sigma}^{-1}(\mathbf{I}_N \otimes \mathbf{\Psi}_v)\mathbf{\Sigma}^{-1} \\
&= \mathbf{I}_N \otimes \mathbf{A}^*\mathbf{\Psi}_v\mathbf{A}^* - (\mathbf{J}_N \otimes \mathbf{A}^*\mathbf{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*) \\
&\quad -(\mathbf{J}_N \otimes \mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{\Psi}_v\mathbf{A}^*) \\
&\quad +N(\mathbf{J}_N \otimes \mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*)
\end{aligned}
$$

with

$$
\begin{aligned}
\frac{\partial l}{\partial \sigma_e^2} &= -\frac{N}{2}\operatorname{tr}\left(\mathbf{A}^*\mathbf{\Psi}_v\right) + \frac{N}{2}\operatorname{tr}\left(\mathbf{A}^*\mathbf{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\right) \\
&\quad +\frac{1}{2}\boldsymbol{\varepsilon}'\left(\mathbf{I}_N \otimes \mathbf{A}^*\mathbf{\Psi}_v\mathbf{A}^*\right)\boldsymbol{\varepsilon} - \widetilde{\boldsymbol{\varepsilon}}'\mathbf{\Psi}_v\overline{\boldsymbol{\varepsilon}} + \frac{N}{2}\overline{\boldsymbol{\varepsilon}}'\mathbf{\Psi}_v\overline{\boldsymbol{\varepsilon}}
\end{aligned}
$$

For σ_u^2 we have

$$
\begin{aligned}
\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial \sigma_u^2}\right) &= \operatorname{tr}(\mathbf{\Sigma}^{-1}(\mathbf{J}_N \otimes \mathbf{\Psi}_\lambda)) = N\operatorname{tr}(\mathbf{A}^*\mathbf{\Psi}_\lambda) - N^2\operatorname{tr}(\mathbf{A}^*\mathbf{\Psi}_\lambda\mathbf{A}^*\mathbf{B}^{-1}), \\
\mathbf{\Sigma}^{-1}\frac{\partial \mathbf{\Sigma}}{\partial \sigma_u^2}\mathbf{\Sigma}^{-1} &= \mathbf{\Sigma}^{-1}(\mathbf{J}_N \otimes \mathbf{\Psi}_\lambda)\mathbf{\Sigma}^{-1} \\
&= (\mathbf{J}_N \otimes \mathbf{A}^*\Psi_\lambda\mathbf{A}^*) \\
&\quad -N(\mathbf{J}_N \otimes \mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\Psi_\lambda\mathbf{A}^*) - N(\mathbf{J}_N \otimes \mathbf{A}^*\Psi_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*) \\
&\quad +N^2(\mathbf{J}_N \otimes \mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\Psi_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial l}{\partial \sigma_u^2} &= -\frac{N}{2}\operatorname{tr}(\mathbf{A}^*\mathbf{\Psi}_\lambda) + \frac{N^2}{2}\operatorname{tr}(\mathbf{A}^*\mathbf{\Psi}_\lambda\mathbf{A}^*\mathbf{B}^{-1}) \\
&\quad +\frac{1}{2}\widetilde{\boldsymbol{\varepsilon}}'\mathbf{\Psi}_\lambda\widetilde{\boldsymbol{\varepsilon}} - N\widetilde{\boldsymbol{\varepsilon}}'\mathbf{\Psi}_\lambda\overline{\boldsymbol{\varepsilon}} + \frac{N^2}{2}\overline{\boldsymbol{\varepsilon}}'\mathbf{\Psi}_\lambda\overline{\boldsymbol{\varepsilon}}
\end{aligned}
$$

Finally for the parameters in $\mathbf{\Psi}_\lambda$ and $\mathbf{\Psi}_v$ let $\mathbf{L}_\lambda = \frac{\partial \mathbf{\Psi}_\lambda}{\partial \rho_\lambda} = \frac{2\rho}{1-\rho^2}\mathbf{\Psi}_\lambda + \frac{1}{1-\rho^2}\mathbf{D}$ where $\mathbf{D}$ is a band matrix with zeros on the main diagonal and $i\rho_\lambda^{i-1}$ on the i$^{\text{th}}$ subdiagonal and define $\mathbf{L}_v$ similarly, we then have

$$\begin{aligned}\frac{\partial l}{\partial \rho_\lambda} &= -\frac{\sigma_u^2 N}{2}\operatorname{tr}(\mathbf{A}^*\mathbf{L}_\lambda) + \frac{\sigma_u^2 N^2}{2}\operatorname{tr}(\mathbf{A}^*\mathbf{L}_\lambda \mathbf{A}^*\mathbf{B}^{-1}) \\ &\quad + \frac{\sigma_u^2}{2}\widetilde{\boldsymbol{\varepsilon}}'\mathbf{L}_\lambda \widetilde{\boldsymbol{\varepsilon}} - \sigma_u^2 N \widetilde{\boldsymbol{\varepsilon}}'\mathbf{L}_\lambda \overline{\boldsymbol{\varepsilon}} + \frac{N^2\sigma_u^2}{2}\overline{\boldsymbol{\varepsilon}}'\mathbf{L}_\lambda \overline{\boldsymbol{\varepsilon}}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial l}{\partial \rho_v} &= -\frac{\sigma_e^2 N}{2}\operatorname{tr}(\mathbf{A}^*\mathbf{L}_v) + \frac{\sigma_e^2 N}{2}\operatorname{tr}(\mathbf{A}^*\mathbf{L}_v \mathbf{A}^*\mathbf{B}^{-1}) \\ &\quad + \frac{\sigma_e^2}{2}\boldsymbol{\varepsilon}'\left(\mathbf{I}_N \otimes \mathbf{A}^*\mathbf{L}_v\mathbf{A}^*\right)\boldsymbol{\varepsilon} - \sigma_e^2 \widetilde{\boldsymbol{\varepsilon}}'\mathbf{L}_v \overline{\boldsymbol{\varepsilon}} + \frac{N\sigma_e^2}{2}\overline{\boldsymbol{\varepsilon}}'\mathbf{L}_v \overline{\boldsymbol{\varepsilon}}\end{aligned}$$

A.2 The information matrix

This appendix derives the elements of the information matrix. For the first element we have the result

$$\mathcal{I}_{\delta,\delta} = \mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z}$$

and the elements $\mathcal{I}_{\delta,\gamma_i}$ are simply computed as

$$\mathcal{I}_{\delta,\gamma_i} = \mathbf{Z}'\boldsymbol{\Sigma}^{-1}\frac{\partial \boldsymbol{\Sigma}}{\partial \gamma_i}\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}$$

Next the elements of the information matrix for the $\boldsymbol{\gamma}$ parameters are obtained as

$$\mathcal{I}_{\gamma_i\gamma_j} = \frac{1}{2}\operatorname{tr}[\boldsymbol{\Sigma}^{-1}(\frac{\partial \boldsymbol{\Sigma}}{\partial \gamma_i})\boldsymbol{\Sigma}^{-1}(\frac{\partial \boldsymbol{\Sigma}}{\partial \gamma_j})]$$

For the elements involving σ_μ^2

$$\mathcal{I}_{\sigma_\mu^2,\sigma_\mu^2} = \frac{1}{2}\left[\begin{array}{c} N\sigma_\alpha^{-4}d^4\left(1-\rho_v\right)^4 - 2N\sigma_\alpha^{-4}d^2\left(1-\rho_v\right)^4 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha} \\ +N^2\sigma_\alpha^{-4}\left(1-\rho_v\right)^4 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}\end{array}\right]$$

where $\mathbf{B}^{-1}$ is defined in appendix A.1.

$$\mathcal{I}_{\sigma_\mu^2,\sigma_e^2} = \frac{1}{2}\left[\begin{array}{c} N\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\boldsymbol{\Psi}_v\boldsymbol{\iota}_T^{\alpha} \\ -2N\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_v\boldsymbol{\iota}_T^{\alpha} \\ +N^2\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}\end{array}\right]$$

$$\mathcal{I}_{\sigma_\mu^2,\sigma_u^2} = \frac{1}{2}\left[\begin{array}{c} N\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\boldsymbol{\iota}_T^{\alpha} \\ -2N^2\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\boldsymbol{\iota}_T^{\alpha} \\ +N^3\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}\end{array}\right]$$

$$\mathcal{I}_{\sigma_\mu^2,\rho_\lambda} = \frac{\sigma_u^2}{2}\left[\begin{array}{c} N\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{L}_\lambda\boldsymbol{\iota}_T^{\alpha} \\ -2N^2\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\boldsymbol{\iota}_T^{\alpha} \\ +N^3\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha} \end{array}\right]$$

$$\mathcal{I}_{\sigma_\mu^2,\rho_v} = \frac{\sigma_e^2}{2}\left[\begin{array}{c} N\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{L}_v\boldsymbol{\iota}_T^{\alpha} \\ -2N\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\boldsymbol{\iota}_T^{\alpha} \\ +N^2\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha} \end{array}\right]$$

with $\mathbf{L}_\lambda$ and $\mathbf{L}_v$ defined in appendix A.1. Next for the relevant $\mathcal{I}_{\sigma_e^2,\gamma_j}$ elements

$$\mathcal{I}_{\sigma_e^2,\sigma_e^2} = \frac{1}{2}\left[\begin{array}{c} N\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\right)^2 - 2N\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_v\right) \\ +N^2\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\right)^2 \end{array}\right]$$

$$\mathcal{I}_{\sigma_e^2,\sigma_u^2} = \frac{1}{2}\left[\begin{array}{c} N\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\boldsymbol{\Psi}_\lambda\right) - 2N^2\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\right) \\ +N^3\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\right) \end{array}\right]$$

$$\mathcal{I}_{\sigma_e^2,\rho_\lambda} = \frac{\sigma_u^2}{2}\left[\begin{array}{c} N\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{L}_\lambda\right) - 2N^2\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\right) \\ +N^3\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\mathbf{A}^*\right) \end{array}\right]$$

$$\mathcal{I}_{\sigma_e^2,\rho_v} = \frac{\sigma_e^2}{2}\left[\begin{array}{c} N\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{L}_v\right) - 2N\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\right) \\ +N^2\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\mathbf{A}^*\right) \end{array}\right]$$

Finally for the elements involving σ_u^2, ρ_λ and ρ_v we have

$$\mathcal{I}_{\sigma_u^2,\sigma_u^2} = \frac{1}{2}\left[\begin{array}{c} N^2\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_\lambda\right)^2 - 2N^3\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\right) \\ +N^4\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\right)^2 \end{array}\right]$$

$$\mathcal{I}_{\sigma_u^2,\rho_\lambda} = \frac{\sigma_u^2}{2}\left[\begin{array}{c} N^2\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\mathbf{L}_\lambda\right) - 2N^3\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\right) \\ +N^4\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\mathbf{A}^*\right) \end{array}\right]$$

$$\mathcal{I}_{\sigma_u^2,\rho_v} = \frac{\sigma_e^2}{2}\left[\begin{array}{c} N\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\mathbf{L}_v\right) - 2N^2\,\mathrm{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\right) \\ +N^3\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\mathbf{A}^*\right) \end{array}\right]$$

$$\mathcal{I}_{\rho_\lambda,\rho_\lambda} = \frac{\sigma_u^4}{2}\left[\begin{array}{c} N^2\,\mathrm{tr}\left(\mathbf{A}^*\mathbf{L}_\lambda\right)^2 - 2N^3\,\mathrm{tr}\left(\mathbf{A}^*\mathbf{L}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\right) \\ +N^4\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\mathbf{A}^*\right)^2 \end{array}\right]$$

$$\mathcal{I}_{\rho_\lambda,\rho_v} = \frac{\sigma_u^2\sigma_e^2}{2}\left[\begin{array}{c} N\,\mathrm{tr}\left(\mathbf{A}^*\mathbf{L}_\lambda\mathbf{A}^*\mathbf{L}_v\right) - 2N^2\,\mathrm{tr}\left(\mathbf{A}^*\mathbf{L}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\right) \\ +N^3\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_\lambda\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\mathbf{A}^*\right) \end{array}\right]$$

$$\mathcal{I}_{\rho_v,\rho_v} = \frac{\sigma_e^4}{2}\left[\begin{array}{c} N\,\mathrm{tr}\left(\mathbf{A}^*\mathbf{L}_v\right)^2 - 2N\,\mathrm{tr}\left(\mathbf{A}^*\mathbf{L}_v\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\right) \\ +N^2\,\mathrm{tr}\left(\mathbf{B}^{-1}\mathbf{A}^*\mathbf{L}_v\mathbf{A}^*\right)^2 \end{array}\right]$$

B Proofs

A number of expressions involving the components of the variance matrix $\boldsymbol{\Sigma}$ appear frequently in the proofs. A series of lemmas below summarizes some basic results for these expressions. Unless otherwise indicated in case of joint convergence $(N, T \to \infty)$ no restriction on the indices are needed and joint limits can also be computed as sequential limits by letting $T \to \infty$ followed by $N \to \infty$, see Phillips and Moon (1999, corollary 1).

Lemma 3 *Let $\mathbf{C}$ be the Prais-Winsten transformation matrix for an $AR(1)$ process with parameter ρ, $\boldsymbol{\Psi}$ the variance covariance matrix of an $AR(1)$ process with parameter r and unit variance and let $\boldsymbol{\iota}_T^{\alpha}$ be a vector with first element $\sqrt{(1+\rho)/(1-\rho)}$ and remaining $T-1$ elements unity.*

$$
\begin{aligned}
\operatorname{tr}\left(\mathbf{C}\boldsymbol{\Psi}\mathbf{C}'\right) &= \frac{2}{1-r^2} - \frac{2(T-1)}{1-r^2} r\rho + \frac{(T-2)}{1-r^2}\left(\rho^2+1\right) \\
\lim_{T\to\infty} \frac{1}{T} \operatorname{tr}\left(\mathbf{C}\boldsymbol{\Psi}\mathbf{C}'\right) &= \frac{\left(\rho^2+1\right) - 2r\rho}{1-r^2}
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{tr}\left(\mathbf{C}\boldsymbol{\Psi}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha}\boldsymbol{\iota}_T^{\alpha\prime}\right) &= \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha} \\
&= \frac{1}{c}\left(\frac{1+\rho}{1-\rho}\right)\left(1-\rho^2\right) \\
&\quad + \frac{1}{c} 2 (r-\rho)\left(1-\rho^2\right)^{1/2} \sqrt{\left(\frac{1+\rho}{1-\rho}\right)} \sum_{j=0}^{T-2} r^j \\
&\quad + \frac{1}{c}(T-1)\left(1-2r\rho+p^2\right) \\
&\quad + \frac{1}{c} 2 \begin{pmatrix} \rho\left(r\rho - r^2\right) \\ +(r-\rho) \end{pmatrix} \sum_{j=2}^{T-1} (T-j)\, r^{j-2} \\
\lim_{T\to\infty} \frac{1}{T} \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha} &= \frac{1}{c}\left(1-2r\rho+p^2\right) - \frac{1}{c(r-1)} 2\left(\rho\left(r\rho - r^2\right) + (r-\rho)\right)
\end{aligned}
$$

where $c = (1-r^2)$. Note that theses matrices are independent of N and that the limits hold when $N, T \to \infty$ as well.

Lemma 4 *Let $\mathbf{A}^* = \mathbf{C}'\left(\sigma_\alpha^{-2}\overline{\mathbf{J}}_T^{\alpha} + \sigma_e^{-2}\overline{\mathbf{E}}_T^{\alpha}\right)\mathbf{C}$ and consider* $\operatorname{vech}\mathbf{A}^*$ *we then have elementwise convergence of* $\operatorname{vech}\mathbf{A}^*$ *to the infinite sequence* $\operatorname{vech}\left(\frac{1}{\sigma_e^2}\boldsymbol{\Psi}_{v,\infty}^{-1}\right)$ *at the rate T^{-1} as $T \to \infty$.*

Let $\mathbf{B}=\mathbf{\Psi}_{\lambda}^{-1}+N\sigma_u^2\mathbf{A}^*$ *we then have*

$$\lim_{T\rightarrow\infty}\operatorname{vech}\mathbf{B}^{-1}=\operatorname{vech}\left(\mathbf{\Psi}_{\lambda,\infty}^{-1}+N\sigma_u^2\sigma_e^{-2}\mathbf{\Psi}_{v,\infty}^{-1}\right)^{-1}$$

$$\begin{aligned}
\lim_{N\rightarrow\infty}\operatorname{vech}\mathbf{B}^{-1} &= \mathbf{0}\\
\lim_{N\rightarrow\infty}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha} &= 0\\
\lim_{T\rightarrow\infty}\frac{1}{T^p}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha} &= 0,\forall p>1\\
\lim_{N\rightarrow\infty}\operatorname{tr}\left(\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{\Psi}\right) &= 0
\end{aligned}$$

$$\begin{aligned}
\lim_{N\rightarrow\infty}\frac{1}{N}\left|\mathbf{B}^{-1}\right| &= 0\\
\lim_{T\rightarrow\infty}\frac{1}{T}\left|\mathbf{B}^{-1}\right| &= 0\\
\lim_{N,T\rightarrow\infty}\frac{1}{NT}\left|\mathbf{B}^{-1}\right| &= 0
\end{aligned}$$

and

$$\begin{aligned}
\lim_{N\rightarrow\infty}\frac{1}{N}\ln\left|\mathbf{B}^{-1}\right| &= 0\\
\lim_{N,T\rightarrow\infty}\frac{1}{NT}\ln\left|\mathbf{B}^{-1}\right| &= 0.
\end{aligned}$$

Proof. To obtain the elementwise convergence of $\operatorname{vech}\mathbf{A}^*$ we write

$$\mathbf{A}^*=\mathbf{C}'\left(\frac{1}{d^2\sigma_\mu^2\left(1-\rho_v\right)^2+\sigma_e^2}\frac{\boldsymbol{\iota}_T^{\alpha}\boldsymbol{\iota}_T^{\alpha\prime}}{d^2}+\sigma_e^{-2}\left(\mathbf{I}_T-\frac{\boldsymbol{\iota}_T^{\alpha}\boldsymbol{\iota}_T^{\alpha\prime}}{d^2}\right)\right)\mathbf{C}$$

and note that $\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}=\left(\sqrt{\frac{(1+\rho_v)(1-\rho_v^2)}{1-\rho_v}}-\rho_v,1-\rho_v,\ldots,1-\rho_v,1\right)$. The elements of $\frac{1}{d^2}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\rightarrow 0$ as $T\rightarrow\infty$ since $d^2=\alpha^2+(T-1)$. Next the established limit for $\operatorname{vech}\mathbf{B}^{-1}$ as $T\rightarrow\infty$ follows from the elementwise convergence of $\operatorname{vech}\mathbf{A}^*$ as $T\rightarrow\infty$ and $\lim_{N\rightarrow\infty}\operatorname{vech}\mathbf{B}^{-1}=\mathbf{0}$ follows since $\lim_{N\rightarrow\infty}\frac{1}{N}\mathbf{B}=\sigma_u^2\mathbf{A}^*$ elementwise. Then $\lim_{N\rightarrow\infty}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}=0$ and $\lim_{N\rightarrow\infty}\operatorname{tr}\left(\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{\Psi}\right)=0$ follows immediately from $\lim_{N\rightarrow\infty}\operatorname{vech}\mathbf{B}^{-1}=\mathbf{0}$. To establish the $T\rightarrow\infty$ limit of $\frac{1}{T^p}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}=0$, $p>1$ we note that for N fix and $T\rightarrow\infty$ $\operatorname{vech}\mathbf{B}^{-1}$ converges elementwise to the infinite sequence $\operatorname{vech}\left(\mathbf{\Psi}_{\lambda,\infty}^{-1}+N\sigma_u^2\sigma_e^{-2}\mathbf{\Psi}_{v,\infty}^{-1}\right)^{-1}$ which has the form of the inverse of a MA(1) covariance matrix, that is the off-diagonal elements decay exponentially. Since $\mathbf{A}^*$ converges elementwise to a band-diagonal matrix it follows that $\mathbf{A}^*\mathbf{B}^{-1}$ converges elementwise to a matrix with exponentially decaying off-diagonals. Hence $\frac{1}{T}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}$ converges to a constant since this is the sum of the exponentially decaying elements in $\mathbf{A}^*\mathbf{B}^{-1}$ and

$$\lim_{T\to\infty}\frac{1}{T^p}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha}=0, p>1$$

follows.

To establish the limits for $|\mathbf{B}^{-1}|$ we note that $\mathbf{B}-\boldsymbol{\Psi}_\lambda^{-1}=N\sigma_u^2\mathbf{A}^*$ is positive definite which implies that $|\mathbf{B}|>\left|\boldsymbol{\Psi}_\lambda^{-1}\right|=1-\rho_\lambda^2$ and $|\mathbf{B}^{-1}|<|\boldsymbol{\Psi}_\lambda|=\frac{1}{1-\rho_\lambda^2}$. In addition $|\mathbf{B}^{-1}|>0$ since $\mathbf{B}$ is positive definite and the results follow.

For $\ln|\mathbf{B}^{-1}|$ we have $\ln|\mathbf{B}^{-1}|<-\ln(1-\rho_\lambda^2)$, a lower bound is obtained from the Hadamard determinant theorem,

$$|\mathbf{B}|\leq\prod_{j=1}^{T}b_{jj}=\prod_{j=1}^{T}\left[\psi^{jj}+N\sigma_u^2a_{jj}^*\right]$$

implying $\ln|\mathbf{B}^{-1}|\geq-\sum_{j=1}^{T}\ln\left[\psi^{jj}+N\sigma_u^2a_{jj}^*\right]\geq-\sum_{j=1}^{T}\ln(1+\rho_\lambda^2+Nk)=-T\ln(1+\rho_\lambda^2+Nk)$ where $k=\max a_{jj}^*$. Note that k depends on T and approaches $\sigma_u^2(1+\rho_\lambda^2)/\sigma_e^2$ as $T\to\infty$ ■

Lemma 5 *Let $\boldsymbol{\Sigma}_0$ be the variance matrix $\boldsymbol{\Sigma}$ evaluated at θ_0. Then*

$$\begin{aligned}\lim_{T\to\infty}\frac{1}{NT}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0 &= \left(1-\frac{1}{N}\right)\frac{\sigma_{e0}^2}{\sigma_e^2}\left[\frac{(\rho_v^2+1)-2\rho_{v0}\rho_v}{1-\rho_{v0}^2}\right]\\ &\quad+\frac{1}{NT}\lim_{T\to\infty}\operatorname{tr}\left(\mathbf{P}_0\mathbf{P}^{-1}\right)\end{aligned}$$

where $\mathbf{P}_0=\left(N\sigma_{u0}^2\boldsymbol{\Psi}_{\lambda0}+\sigma_{e0}^2\boldsymbol{\Psi}_{v0}\right)$, $\mathbf{P}=\left(N\sigma_u^2\boldsymbol{\Psi}_\lambda+\sigma_e^2\boldsymbol{\Psi}_v\right)$.

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{NT}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0 &= \frac{\sigma_{\mu0}^2}{T\sigma_\alpha^2}d^2(1-\rho_v)^2+\frac{\sigma_{e0}^2}{T\sigma_e^2}\operatorname{tr}\left(\mathbf{C}\boldsymbol{\Psi}_{v0}\mathbf{C}'\right)\\ &\quad+\frac{\sigma_{e0}^2\left(\sigma_\alpha^{-2}-\sigma_e^{-2}\right)}{Td^2}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}_{v0}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha}\end{aligned}$$

with $\operatorname{tr}\left(\mathbf{C}\boldsymbol{\Psi}_{v0}\mathbf{C}'\right)$ and $\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}_{v0}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha}$ evaluated in lemma 3, and

$$\lim_{N,T\to\infty}\frac{1}{NT}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0=\frac{\sigma_{e0}^2}{\sigma_e^2}\left[\frac{(\rho_v^2+1)-2\rho_{v0}\rho_v}{1-\rho_{v0}^2}\right]$$

Proof. Standard matrix algebra yields

$$\begin{aligned}
\frac{1}{NT}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0 &= \frac{1}{T}\sigma_{\mu0}^2\sigma_\alpha^{-2}d^2\left(1-\rho_v\right)^2 \\
&+\frac{\sigma_{e0}^2\left(\sigma_\alpha^{-2}-\sigma_e^{-2}\right)}{Td^2}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}_{v0}\mathbf{C}'\boldsymbol{\iota}_T^\alpha \\
&+\frac{\sigma_{u0}^2\left(\sigma_\alpha^{-2}-\sigma_e^{-2}\right)}{Td^2}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}_{\lambda0}\mathbf{C}'\boldsymbol{\iota}_T^\alpha \\
&+\frac{1}{T}\sigma_{e0}^2\sigma_e^{-2}\operatorname{tr}\left(\mathbf{C}\boldsymbol{\Psi}_{v0}\mathbf{C}'\right)+\frac{1}{T}\sigma_{u0}^2\sigma_e^{-2}\operatorname{tr}\left(\mathbf{C}\boldsymbol{\Psi}_{\lambda0}\mathbf{C}'\right) \\
&-\frac{1}{T}\sigma_{\mu0}^2\sigma_u^2\sigma_\alpha^{-2}\left(1-\rho_v\right)^2\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^\alpha \\
&-\frac{1}{T}\sigma_{e0}^2\sigma_u^2\operatorname{tr}(\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_{v0}) \\
&-\frac{1}{T}N\sigma_{u0}^2\sigma_u^2\operatorname{tr}(\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_{\lambda0})
\end{aligned}$$

To establish the limit as $T\to\infty$ note that for the first term

$$\lim_{T\to\infty}\frac{1}{T}\sigma_{\mu0}^2\sigma_\alpha^{-2}d^2\left(1-\rho_v\right)^2=0$$

since $\sigma_{\mu0}^2\sigma_\alpha^{-2}d^2\left(1-\rho_v\right)^2=O\left(1\right)$. For the next two terms

$$\begin{aligned}
\lim_{T\to\infty}\frac{1}{T}\sigma_\alpha^{-2}d^{-2}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}_{j0}\mathbf{C}'\boldsymbol{\iota}_T^\alpha &= 0\ ,j=\lambda,v \\
\lim_{T\to\infty}\frac{1}{T}d^{-2}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}_{j0}\mathbf{C}'\boldsymbol{\iota}_T^\alpha &= 0\ ,j=\lambda,v
\end{aligned}$$

follows from lemma 3 since $\sigma_\alpha^2=O\left(T\right)$ and $d^2=O\left(T\right)$. The limits as $T\to\infty$ of the fourth and fifth term follow from lemma 3 and lemma 4 established that the sixth term converges to zero. For the last two terms we have by the elementwise convergence of $\operatorname{vech}\mathbf{A}^*$ and $\operatorname{vech}\mathbf{B}^{-1}$ established in lemma 4

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T}\operatorname{tr}\left(\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\boldsymbol{\Psi}_{j0}\right) \\
=\ &\lim_{T\to\infty}\frac{1}{T}\operatorname{tr}\left(\sigma_e^{-2}\boldsymbol{\Psi}_v^{-1}\left(\boldsymbol{\Psi}_\lambda^{-1}+N\sigma_u^2\sigma_e^{-2}\boldsymbol{\Psi}_v^{-1}\right)^{-1}\sigma_e^{-2}\boldsymbol{\Psi}_v^{-1}\boldsymbol{\Psi}_{j0}\right)
\end{aligned}$$

which is well defined (and non-zero) since the diagonal elements of the matrix are $O\left(1\right)$. Repeatedly applying elementary results on inverses of sums (Dhrymes (1984, p. 39)) to these last two terms and then collecting terms obtains the expression given in the theorem. This completes the proof of the $T\to\infty$ case.

Now consider the case when $N \to \infty$. Since all but the three last terms are independent of N we need only consider these. Then

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{T}\sigma_{\mu0}^2\sigma_u^2\sigma_\alpha^{-2}\left(1-\rho_v\right)^2 \boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{B}^{-1}\boldsymbol{\iota}_T^{\alpha} &= 0 \\ \lim_{N\to\infty} \frac{1}{T}\sigma_{u0}^2\sigma_u^2 \operatorname{tr}(\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{\Psi}_{\lambda 0}) &= 0 \end{aligned}$$

follows from lemma 3, and

$$\lim_{N\to\infty} \frac{1}{T}N\sigma_{u0}^2\sigma_u^2 \operatorname{tr}(\mathbf{A}^*\mathbf{B}^{-1}\mathbf{A}^*\mathbf{\Psi}_{\lambda 0}) = \frac{1}{T}\sigma_{u0}^2 \operatorname{tr}(\mathbf{\Psi}_{\lambda 0}\mathbf{A}^*)$$

follows since $\lim_{N\to\infty} N\mathbf{B}^{-1} = \left(\sigma_u^2\mathbf{A}^*\right)^{-1}$ elementwise. Collecting terms as in the $T \to \infty$ case then gives the result.

Finally the result for $N, T \to \infty$ follows by taking sequential limits and using lemma 3 ■

The following lemma gives some basic limit results for the expressions $\mathbf{Z}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon}$, $\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\iota}_{NT}, \mathbf{D}'\mathbf{\Sigma}^{-1}\mathbf{D}$ and $\mathbf{H}'\mathbf{\Sigma}^{-1}\mathbf{H}$. In the proof of the results in this lemma we make extensive use of elementary results on inverses involving sums (Dhrymes (1984, p. 39)), applying them repeatedly to obtain manageable expressions.

Lemma 6 *As $N \to \infty, T \to \infty$ or $N, T \to \infty$*

$$\operatorname{plim} \frac{1}{NT}\mathbf{X}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon} = \mathbf{0}$$

If both $N, T \to \infty$ and if $\frac{N}{T} \to \infty$

$$\begin{aligned} \operatorname*{plim}_{N,T\to\infty} \frac{1}{T}\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon} &= 0 \\ \lim_{N,T\to\infty} \frac{1}{T}\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\iota}_{NT} &= \frac{1}{\sigma_u^2}\left(1-\rho_\lambda\right)^2 \end{aligned}$$

If both $N, T \to \infty$ and if $\frac{T}{N} \to \infty$

$$\begin{aligned} \operatorname*{plim}_{N,T\to\infty} \frac{1}{N}\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon} &= 0 \\ \lim_{N,T\to\infty} \frac{1}{N}\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\iota}_{NT} &= \frac{1}{\sigma_\mu^2} \end{aligned}$$

If $N, T \to \infty$ simultaneously

$$\begin{aligned} \operatorname*{plim}_{N,T\to\infty} \frac{1}{N}\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon} &= \operatorname*{plim}_{N,T\to\infty} \frac{1}{T}\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon} = 0 \\ \lim_{N,T\to\infty} \frac{1}{N}\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\iota}_{NT} &= \lim_{N,T\to\infty} \frac{1}{T}\boldsymbol{\iota}_{NT}'\mathbf{\Sigma}^{-1}\boldsymbol{\iota}_{NT} = \frac{1}{\sigma_\mu^2 + \sigma_u^2\left(1-\rho_\lambda\right)^{-2}} \end{aligned}$$

As $T \to \infty$ *or* $N, T \to \infty$

$$\text{plim}\,\frac{1}{T}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} = \mathbf{0}$$

As $N \to \infty$

$$\text{plim}\,\frac{1}{NT}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} = \mathbf{0}$$

As $N \to \infty$ *or* $N, T \to \infty$

$$\text{plim}\,\frac{1}{N}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} = \mathbf{0}$$

As $T \to \infty$

$$\text{plim}\,\frac{1}{NT}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} = \mathbf{0}$$

If both $N, T \to \infty$ *and if* $\frac{N}{T} \to \infty$

$$\begin{aligned}
\underset{N,T\to\infty}{\text{plim}}\,\frac{1}{T}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D} &= \frac{1}{T\sigma_{u0}^2}\underset{T\to\infty}{\text{plim}}\sum_{t=2}^{T}\left(\mathbf{d}_t - \rho_\lambda \mathbf{d}_{t-1}\right)\left(\mathbf{d}_t - \rho_\lambda \mathbf{d}_{t-1}\right)' \\
\underset{N,T\to\infty}{\text{plim}}\,\frac{1}{N}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H} &= \frac{1}{\sigma_\mu^2}\underset{N\to\infty}{\text{plim}}\,\frac{1}{N}\mathbf{h}'\overline{\mathbf{E}}_N\mathbf{h}
\end{aligned}$$

where $\overline{\mathbf{E}}_N = \mathbf{I}_N - \overline{\mathbf{J}}_N, \overline{\mathbf{J}}_N = \frac{1}{N}\boldsymbol{\iota}_N\boldsymbol{\iota}_N'$. *If both* $N, T \to \infty$ *and if* $\frac{T}{N} \to \infty$

$$\begin{aligned}
\underset{N,T\to\infty}{\text{plim}}\,\frac{1}{T}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D} &= \underset{T\to\infty}{\text{plim}}\,\frac{1}{T\sigma_u^2}\mathbf{S}_\lambda \\
\underset{N,T\to\infty}{\text{plim}}\,\frac{1}{N}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H} &= \frac{1}{\sigma_\mu^2}\underset{N\to\infty}{\text{plim}}\,\frac{1}{N}\mathbf{h}'\mathbf{h}
\end{aligned}$$

where $\mathbf{S}_\lambda = \left(\sum_{t=2}^{T}\left(\mathbf{d}_t - \rho_\lambda \mathbf{d}_{t-1}\right)\left(\mathbf{d}_t - \rho_\lambda \mathbf{d}_{t-1}\right)' - \frac{(1-\rho_{\lambda 0})^2}{T}\sum_{t=2}^{T-1}\sum_{r=2}^{T-1}\mathbf{d}_t\mathbf{d}_r'\right)$.
Finally, if $N, T \to \infty$ *simultaneously*

$$\begin{aligned}
\underset{N,T\to\infty}{\text{plim}}\,\frac{1}{T}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D} &= \underset{T\to\infty}{\text{plim}}\,\frac{1}{T\sigma_u^2}\mathbf{S}_\lambda + \underset{T\to\infty}{\text{plim}}\,\frac{1}{T^2\left(\sigma_\mu^2 + \frac{\sigma_u^2}{(1-\rho_\lambda)^2}\right)}\sum_{t=2}^{T-1}\sum_{r=2}^{T-1}\mathbf{d}_t\mathbf{d}_r' \\
\underset{N,T\to\infty}{\text{plim}}\,\frac{1}{N}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H} &= \underset{N\to\infty}{\text{plim}}\,\frac{1}{N\sigma_\mu^2}\mathbf{h}'\overline{\mathbf{E}}_N\mathbf{h} + \underset{N\to\infty}{\text{plim}}\,\frac{1}{N\left(\sigma_\mu^2 + \frac{\sigma_u^2}{(1-\rho_\lambda)^2}\right)}\mathbf{h}'\overline{\mathbf{J}}_N\mathbf{h}
\end{aligned}$$

Proof. To obtain the limit results for $\frac{1}{NT}\mathbf{X}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon}$ we write

$$\begin{aligned}
\frac{1}{NT}\mathbf{X}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon} &= \frac{(1-\rho_v)}{NT\sigma_\alpha^2}\sum_{t=1}^{T}\sum_{i=1}^{N}X_{i,t}r_t\mu_i \\
&-\frac{(1-\rho_v)}{N^2T\sigma_\alpha^2}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}X_{i,t}r_t\mu_j+\frac{1}{N^2T}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}X_{i,t}c_t\mu_j \\
&+\frac{1}{NT}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{i=1}^{N}X_{i,t}A^*_{t,r}v_{i,r} \\
&-\frac{1}{N^2T}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}X_{i,t}A^*_{t,r}v_{j,r} \\
&+\frac{1}{N^2T}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}X_{i,t}L^{t,r}v_{j,r} \\
&+\frac{1}{NT}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{i=1}^{N}X_{i,t}L^{t,r}\lambda_r
\end{aligned} \tag{7}$$

where r_t denotes the $t:th$ element of $\mathbf{C}'\boldsymbol{\iota}_T^\alpha$ and c_t denotes an element of the $T\times 1$ vector $\left(N\sigma_u^2\mathbf{\Psi}_\lambda+(\mathbf{A}^*)^{-1}\right)^{-1}\boldsymbol{\iota}_T$. Further $L^{t,r}$ denotes the $tr:th$ element of the $T\times T$ matrix $\mathbf{L}^{-1}=\left(N\sigma_u^2\mathbf{\Psi}_\lambda+(\mathbf{A}^*)^{-1}\right)^{-1}$ and $A^*_{t,r}$ denotes the $tr:th$ element of the $T\times T$ matrix $\mathbf{A}^*$.

First we consider the probability limits of the terms involving μ_i

$$\begin{aligned}
\operatorname{plim}\frac{(1-\rho_v)}{NT\sigma_\alpha^2}\sum_{t=1}^{T}\sum_{i=1}^{N}X_{i,t}r_t\mu_i &= 0 \\
\operatorname{plim}\frac{(1-\rho_v)}{N^2T\sigma_\alpha^2}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}X_{i,t}r_t\mu_j &= 0
\end{aligned}$$

as $N\to\infty, T\to\infty$ or $N,T\to\infty$ are straightforward to show since $\mathbf{C}'\boldsymbol{\iota}_T^\alpha$ is a constant vector. To establish corresponding results for

$$\operatorname{plim}\frac{1}{N^2T}\sum_{t=1}^{T}\sum_{j=1}^{N}\sum_{i=1}^{N}X_{i,t}c_t\mu_j \tag{8}$$

we need to consider the properties of c_t. First, since c_t is the $t:th$ element of $\mathbf{L}\boldsymbol{\iota}_T=\sigma_\alpha^{-2}(1-\rho_v)\mathbf{B}^{-1}\mathbf{C}'\boldsymbol{\iota}_T^\alpha$ and $\mathbf{B}^{-1}$ converges elementwise to a matrix with exponentially decaying off-diagonals, every element of $\mathbf{L}\boldsymbol{\iota}_T$ is an exponentially decaying sum. Secondly, by the properties of $\mathbf{B}^{-1}$ every element of $\mathbf{L}\boldsymbol{\iota}_T$ is $O\left((NT)^{-1}\right)$. This shows that (8) is zero as either or both of the indices grows large.

For the elements involving the idiosyncratic errors, v_{it}

$$\begin{aligned} \operatorname{plim} \frac{1}{NT} \sum_{r=1}^{T} \sum_{t=1}^{T} \sum_{i=1}^{N} X_{i,t} A^*_{t,r} v_{i,r} &= 0 \\ \operatorname{plim} \frac{1}{N^2T} \sum_{r=1}^{T} \sum_{t=1}^{T} \sum_{j=1}^{N} \sum_{i=1}^{N} X_{i,t} A^*_{t,r} v_{j,r} &= 0 \end{aligned}$$

as $N \to \infty, T \to \infty$ or $N, T \to \infty$ holds since $\mathbf{A}^* = \mathbf{C}'\left(\sigma_\alpha^{-2}\overline{\mathbf{J}}_T^\alpha + \sigma_e^{-2}\overline{\mathbf{E}}_T^\alpha\right)\mathbf{C}$, $\mathbf{C}'\mathbf{C}$ is band-diagonal and $\mathbf{C}'\boldsymbol{\iota}_T^\alpha$ is a constant vector. To be able to write

$$\operatorname{plim} \frac{1}{N^2T} \sum_{r=1}^{T} \sum_{t=1}^{T} \sum_{j=1}^{N} \sum_{i=1}^{N} X_{i,t} L^{t,r} v_{j,r} = 0 \tag{9}$$

as $N \to \infty, T \to \infty$ or $N, T \to \infty$ we need to establish some properties of $\mathbf{L}^{-1}$. For this purpose we let $\mathbf{C}\boldsymbol{\Psi}_v\mathbf{C}' = \mathbf{I}_T, \mathbf{Q}$ be the eigenvectors of $\boldsymbol{\Psi}_\lambda$ in the metric of $\boldsymbol{\Psi}_v$. That is, $\mathbf{C}\boldsymbol{\Psi}_v\mathbf{C}' = \mathbf{Q}\boldsymbol{\Lambda}$ where $\boldsymbol{\Lambda}$ is diagonal and $\mathbf{Q}$ is orthogonal. Further let $\mathbf{W} = \mathbf{C}'\mathbf{Q}$ we can then write

$$\begin{aligned} \boldsymbol{\Psi}_v &= \mathbf{W}^{-1\prime}\mathbf{W}^{-1} = \mathbf{C}^{-1}\mathbf{Q}\mathbf{Q}'\mathbf{C}^{-1\prime} \\ \boldsymbol{\Psi}_\lambda &= \mathbf{C}^{-1}\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}'\mathbf{C}^{-1\prime} = \mathbf{W}^{-1\prime}\boldsymbol{\Lambda}\mathbf{W}^{-1} \end{aligned}$$

and

$$\begin{aligned} \mathbf{L} &= \mathbf{W}^{-1\prime}\mathbf{W}'\left(N\sigma_u^2\boldsymbol{\Psi}_\lambda + \sigma_\mu^2\mathbf{J}_T + \sigma_e^2\boldsymbol{\Psi}_v\right)\mathbf{W}\mathbf{W}^{-1} \\ &= \mathbf{W}^{-1\prime}\left(\mathbf{D} + \left(\sigma_\alpha^2 - \sigma_e^2\right)\overline{\mathbf{J}}_T^w\right)\mathbf{W}^{-1} \end{aligned}$$

where $\overline{\mathbf{J}}_T^w = \mathbf{Q}'\overline{\mathbf{J}}_T^\alpha\mathbf{Q}$ is idempotent, $\mathbf{D} = N\sigma_u^2\boldsymbol{\Lambda} + \sigma_e^2\mathbf{I}_T$ is diagonal. Since $\boldsymbol{\Lambda}$ is diagonal with bounded constant elements setting $\boldsymbol{\Lambda} = \varphi\mathbf{I}_T$ will not change the order properties of $\mathbf{L}^{-1}$. Hence, defining $\overline{\mathbf{E}}_T^w = \mathbf{I}_T - \overline{\mathbf{J}}_T^w$ we obtain

$$\mathbf{L}^{-1} \approx \mathbf{W}'\left(\frac{1}{(N\sigma_u^2\varphi + \sigma_e^2)}\overline{\mathbf{E}}_T^w + \frac{1}{(N\sigma_u^2\varphi + \sigma_\alpha^2)}\overline{\mathbf{J}}_T^w\right)\mathbf{W}$$

which shows that $\mathbf{L}^{-1}$ is similar to $\mathbf{A}^*$ except that the elements of $\mathbf{L}^{-1}$ are $O\left(N^{-1}\right)$. This shows that (9) holds.

Finally for the term involving λ_t

$$\operatorname*{plim}_{N\to\infty} \frac{1}{NT} \sum_{r=1}^{T} \sum_{t=1}^{T} \sum_{i=1}^{N} X_{i,t} L^{t,r} \lambda_r = 0$$

follows since the elements of $\mathbf{L}^{-1}$ are $O\left(N^{-1}\right)$. Next

$$\operatorname*{plim}_{T\to\infty}\frac{1}{NT}\sum_{r=1}^{T}\sum_{t=1}^{T}\sum_{i=1}^{N}X_{i,t}L^{t,r}\lambda_r=0$$

since λ_t have zero mean and by the properties of $\mathbf{L}^{-1}$. It follows that the probability limit is zero as both N and $T\to\infty$ as well. This completes the proof of the first result in the lemma. We consider next the limits of the terms involving the constant.

To obtain results for $\boldsymbol{\iota}_{NT}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}$ we let $X_{it}=1\ \forall i,t$ in (7). This gives

$$\boldsymbol{\iota}_{NT}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}=(1-\rho_v)\sum_{t=1}^{T}\sum_{i=1}^{N}c_t\mu_i+\sum_{t=1}^{T}\sum_{i=1}^{N}c_t v_{i,t}+N\sum_{t=1}^{T}c_t\lambda_t$$

If $\frac{N}{T}\to\infty$ we normalize by $\frac{1}{T}$ to obtain

$$\begin{aligned}\operatorname*{plim}_{N\to\infty}\frac{(1-\rho_v)}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}c_t\mu_i &= 0\\ \operatorname*{plim}_{N\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N}c_t v_{i,t} &= 0\\ \operatorname*{plim}_{T\to\infty}\operatorname*{plim}_{N\to\infty}\frac{N}{T}\sum_{t=1}^{T}c_t\lambda_t &= 0\end{aligned}$$

as a consequence of the properties of c_t. To obtain corresponding results for $\boldsymbol{\iota}_{NT}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota}_{NT}$ write

$$\begin{aligned}&\frac{1}{NT}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}\\ =\ &\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{A}^*\mathbf{X}_i-\frac{1}{N^2T}\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{X}_j'\mathbf{A}^*\mathbf{X}_i+\frac{1}{N^2T}\sum_{j=1}^{N}\sum_{i=1}^{N}\mathbf{X}_j'\mathbf{L}^{-1}\mathbf{X}_i\end{aligned}$$

and note that if $\mathbf{X}_i=\boldsymbol{\iota}_T\ \forall i$ we arrive at

$$\begin{aligned}\boldsymbol{\iota}_{NT}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota}_{NT} &= N\boldsymbol{\iota}_T'\mathbf{L}^{-1}\boldsymbol{\iota}_T=N\boldsymbol{\iota}_T'\left(N\sigma_u^2\boldsymbol{\Psi}_\lambda+(\mathbf{A}^*)^{-1}\right)^{-1}\boldsymbol{\iota}_T\\ &= \sigma_\alpha^{-2}(1-\rho_v)N\boldsymbol{\iota}_T'\left(N\sigma_u^2\boldsymbol{\Psi}_\lambda\mathbf{A}^*+\mathbf{I}_T\right)^{-1}\mathbf{C}'\boldsymbol{\iota}_T^\alpha\\ &= \sigma_\alpha^{-2}\sigma_u^{-2}(1-\rho_v)\boldsymbol{\iota}_T'\left(\boldsymbol{\Psi}_\lambda\mathbf{A}^*\right)^{-1}\mathbf{C}'\boldsymbol{\iota}_T^\alpha+O\left(N^{-1}\right)\\ &= \sigma_\alpha^{-2}\sigma_u^{-2}(1-\rho_v)\boldsymbol{\iota}_T'\left(\mathbf{A}^*\right)^{-1}\boldsymbol{\Psi}_\lambda^{-1}\mathbf{C}'\boldsymbol{\iota}_T^\alpha+O\left(N^{-1}\right)\end{aligned}$$

since $\eta_t = \boldsymbol{\iota}_T' \mathbf{C}^{-1} \boldsymbol{\iota}_T^{\alpha} = \frac{\alpha}{\sqrt{1-\rho_v^2}} \sum_{j=0}^{T} \rho_v^j + \sum_{j=0}^{T} (T-j) \rho_v^j = O(T)$

$$\begin{aligned}
& \boldsymbol{\iota}_{NT}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\iota}_{NT} \\
= \; & \sigma_{\alpha}^{-2} \sigma_u^{-2} (1-\rho_v) \boldsymbol{\iota}_T' \mathbf{C}^{-1} \left(\frac{\sigma_{\alpha}^2 - \sigma_e^2}{d^2} \boldsymbol{\iota}_T^{\alpha} \boldsymbol{\iota}_T^{\alpha\prime} + \sigma_e^2 \mathbf{I}_T \right) \mathbf{C}^{-1\prime} \boldsymbol{\Psi}_{\lambda}^{-1} \mathbf{C}' \boldsymbol{\iota}_T^{\alpha} + O\left(N^{-1}\right) \\
= \; & \sigma_{\alpha}^{-2} \sigma_u^{-2} (1-\rho_v) \left(\sigma_{\mu}^2 (1-\rho_v)^2 \eta_t \boldsymbol{\iota}_T^{\alpha\prime} + \sigma_e^2 \boldsymbol{\iota}_T' \mathbf{C}^{-1} \right) \mathbf{C}^{-1\prime} \boldsymbol{\Psi}_{\lambda}^{-1} \mathbf{C}' \boldsymbol{\iota}_T^{\alpha} + O\left(N^{-1}\right)
\end{aligned}$$

hence if both $N, T \to \infty$ and $\frac{N}{T} \to \infty$

$$\begin{aligned}
\lim_{T\to\infty} \lim_{N\to\infty} \frac{1}{T} \boldsymbol{\iota}_{NT}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\iota}_{NT} &= \lim_{T\to\infty} \frac{\eta_t \sigma_{\mu}^2}{T \sigma_{\alpha}^2 \sigma_u^2} (1-\rho_v)^3 \boldsymbol{\iota}_T^{\alpha\prime} \mathbf{C}^{-1\prime} \boldsymbol{\Psi}_{\lambda}^{-1} \mathbf{C}' \boldsymbol{\iota}_T^{\alpha} \\
&= \lim_{T\to\infty} \frac{\eta_t \sigma_{\mu}^2}{T \sigma_{\alpha}^2 \sigma_u^2} (1-\rho_v)^2 \boldsymbol{\iota}_T' \mathbf{C}' \mathbf{C}^{-1\prime} \boldsymbol{\Psi}_{\lambda}^{-1} \mathbf{C}' \boldsymbol{\iota}_T^{\alpha} \\
&= \lim_{T\to\infty} \frac{\eta_t}{T d^2 \sigma_u^2} (1-\rho_{\lambda}) \boldsymbol{\iota}_T^{\beta\prime} \mathbf{C}_{\lambda} \mathbf{C}' \boldsymbol{\iota}_T^{\alpha} \\
&= \frac{1}{\sigma_u^2} (1-\rho_{\lambda})^2
\end{aligned}$$

where $\boldsymbol{\iota}_T^{\beta\prime} = \boldsymbol{\iota}_T' \mathbf{C}_{\lambda}$ and $\mathbf{C}_{\lambda}$ is the Prais-Winsten transformation matrix for $\boldsymbol{\Psi}_{\lambda}$. Alternatively this can be derived by noting that $\lim_{N\to\infty} \operatorname{vech}\left(\frac{1}{N}\mathbf{L}\right) = \operatorname{vech}\left(\sigma_u^2 \boldsymbol{\Psi}_{\lambda}\right)$ and hence $\lim_{N\to\infty} \frac{N}{T} \boldsymbol{\iota}_T' \mathbf{L}^{-1} \boldsymbol{\iota}_T = \frac{(1-\rho_{\lambda})^2}{T\sigma_u^2} \boldsymbol{\iota}_T^{\beta\prime} \boldsymbol{\iota}_T^{\beta} \to \frac{(1-\rho_{\lambda})^2}{\sigma_u^2}$ as $T \to \infty$. If $\frac{T}{N} \to \infty$ we have

$$\begin{aligned}
\operatorname*{plim}_{N\to\infty} \operatorname*{plim}_{T\to\infty} \frac{(1-\rho_v)}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} c_t \mu_i &= 0 \\
\operatorname*{plim}_{T\to\infty} \frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} c_t v_{it} &= 0 \\
\operatorname*{plim}_{T\to\infty} \sum_{t=1}^{T} c_t \lambda_t &= 0
\end{aligned}$$

For $\boldsymbol{\iota}_{NT}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\iota}_{NT}$ we arrive at

$$\lim_{T\to\infty} \boldsymbol{\iota}_{NT}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\iota}_{NT} = \lim_{T\to\infty} N \boldsymbol{\iota}_T' \left(N\sigma_u^2 \boldsymbol{\Psi}_{\lambda} + \sigma_{\mu}^2 \mathbf{J}_T + \sigma_e^2 \boldsymbol{\Psi}_v \right)^{-1} \boldsymbol{\iota}_T$$

and hence proceeding by induction

$$\lim_{N\to\infty} \lim_{T\to\infty} \frac{1}{N} \boldsymbol{\iota}_{NT}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\iota}_{NT} = \frac{1}{\sigma_{\mu}^2}$$

Finally if $N,T \to \infty$ simultaneously we obtain

$$\begin{aligned}
\operatorname*{plim}_{N,T\to\infty} \frac{1}{N}\boldsymbol{\iota}'_{NT}\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} &= \operatorname*{plim}_{N,T\to\infty} \frac{1}{T}\boldsymbol{\iota}'_{NT}\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} = 0 \\
\lim_{N,T\to\infty} \frac{1}{N}\boldsymbol{\iota}'_{NT}\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota}_{NT} &= \lim_{N,T\to\infty} \frac{1}{T}\boldsymbol{\iota}'_{NT}\boldsymbol{\Sigma}^{-1}\boldsymbol{\iota}_{NT} = \frac{1}{\sigma_\mu^2 + \sigma_u^2 (1-\rho_\lambda)^{-2}}
\end{aligned}$$

This completes the proof for the terms involving the constant and we proceed to consider the limit results for the terms involving time-invariant explanatory variables or individual-invariant explanatory variables.

To prove that

$$\operatorname{plim} \frac{1}{T}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} = \operatorname{plim} \frac{1}{T}\sum_{i=1}^{N} \mathbf{d}'\mathbf{L}^{-1}\boldsymbol{\iota}_T\mu_i + \operatorname{plim} \frac{1}{T}\sum_{i=1}^{N} \mathbf{d}'\mathbf{L}^{-1}\mathbf{v}_i + \operatorname{plim} \frac{N}{T}\mathbf{d}'\mathbf{L}^{-1}\boldsymbol{\lambda}$$

is a null vector as either $T \to \infty$ or $N,T \to \infty$ it suffices to note the properties of $\mathbf{L}^{-1}$. By the properties of $\mathbf{L}^{-1}$ we similarly have $\operatorname{plim}_{N\to\infty} \frac{N}{T}\mathbf{d}'\mathbf{L}^{-1}\boldsymbol{\lambda} \neq \mathbf{0}$ and $\operatorname{plim}_{N\to\infty} \frac{1}{T}\mathbf{d}'\mathbf{L}^{-1}\boldsymbol{\lambda} = \mathbf{0}$. Results for $\operatorname{plim} \frac{1}{N}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}$ and $\operatorname{plim} \frac{1}{NT}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}$ can be shown analogously. Remaining results can be derived by noting that $\frac{1}{T}\mathbf{D}'\boldsymbol{\Sigma}^{-1}\mathbf{D} = \frac{N}{T}\mathbf{d}'\mathbf{L}^{-1}\mathbf{d}$, and $\frac{1}{N}\mathbf{H}'\boldsymbol{\Sigma}^{-1}\mathbf{H} = \frac{d^2(1-\rho_v)^2}{N\sigma_\alpha^2}\mathbf{h}'\overline{\mathbf{E}}_N\mathbf{h} + \frac{\boldsymbol{\iota}'_\mathbf{T}\mathbf{L}^{-1}\boldsymbol{\iota}_\mathbf{T}}{N}\mathbf{h}'\overline{\mathbf{J}}_N\mathbf{h}$ ■

The next and final lemma gives some important results about the limit behavior of the information cross-elements for the mean parameters, $\boldsymbol{\delta}$ and the elements $\mathcal{I}_{\delta,\gamma}$. Limit results for the elements $\mathcal{I}_{\gamma,\gamma}$ appear in theorem 2. To summarize some of the content in this lemma we can say that the set of consistent parameters (as $N \to \infty$ or $T \to \infty$) are information block-diagonal to the set of inconsistent parameters and that the set of consistent mean parameters and the set of consistent variance parameters are always information block-diagonal.

Lemma 7 *As either or both of N and $T \to \infty$ the cross-elements (properly normalized of course) $\mathcal{I}_{\beta,(\gamma,\alpha)}$, $\mathcal{I}_{\pi,(\gamma_1,\gamma_2)}$, $\mathcal{I}_{\tau,(\gamma_2,\gamma_3)}$ and $\mathcal{I}_{\alpha,\gamma_2}$, converge elementwise to zero in probability (or in expectation), where $\gamma_1 = \sigma_\mu^2$, $\gamma_2 = (\sigma_e^2, \rho_v)$ and $\gamma_3 = (\sigma_u^2, \rho_\lambda)$. As $N \to \infty$ (no matter what T is) this holds for $\mathcal{I}_{\gamma_1,(\alpha,\tau,\pi,\gamma_2)}$, $\mathcal{I}_{\gamma_2,\gamma_3}$, $\mathcal{I}_{\beta,\pi}$ and as $T \to \infty$ (no matter what N is) for $\mathcal{I}_{\gamma_2,(\alpha,\pi,\tau,\gamma_1)}$, $\mathcal{I}_{\gamma_2,\gamma_3}$, and $\mathcal{I}_{\beta,\tau}$. We now concentrate on mainly the non-zero cross elements of interest. If only $N \to \infty$*

$$\operatorname{plim} \frac{1}{\sqrt{T}}\mathcal{I}_{\alpha,\pi} = \frac{(1-\rho_\lambda)}{\sigma_u^2\sqrt{T}}\mathbf{d}'\mathbf{C}'_\lambda\boldsymbol{\iota}_T^\beta$$

where $\mathbf{C}'_\lambda\mathbf{C}_\lambda = \boldsymbol{\Psi}_\lambda^{-1}, (1-\rho_\lambda)\,\boldsymbol{\iota}_T^\beta = \mathbf{C}_\lambda\boldsymbol{\iota}_T$, hence if $N,T \to \infty$ such that $\frac{N}{T} \to \infty$

$$\operatorname{plim} \frac{1}{T}\mathcal{I}_{\alpha,\pi} = \frac{(1-\rho_\lambda)^2}{\sigma_u^2} \operatorname*{plim}_{T\to\infty} \frac{1}{T}\sum_{t=2}^{T-1} \mathbf{d}'_t$$

and if $T \to \infty$ or $N, T \to \infty$ such that $\frac{T}{N} \to \infty$

$$\text{plim}\, \frac{1}{\sqrt{T}} \mathcal{I}_{\alpha,\pi} = \mathbf{0}$$

If $N, T \to \infty$ simultaneously

$$\text{plim}\, \frac{1}{T} \mathcal{I}_{\alpha,\pi} = \frac{1}{\sigma_u^2 \left(1-\rho_\lambda\right)^{-2} + \sigma_\mu^2} \underset{T\to\infty}{\text{plim}}\, \frac{1}{T} \sum_{t=2}^{T-1} \mathbf{d}_t'$$

If only $T \to \infty$

$$\text{plim}\, \frac{1}{\sqrt{N}} \mathcal{I}_{\alpha,\tau} = \frac{1}{\sqrt{N}\sigma_\mu^2} \sum_{i=1}^{N} \mathbf{h}_i'$$

hence if $N, T \to \infty$ such that $\frac{T}{N} \to \infty$

$$\text{plim}\, \frac{1}{N} \mathcal{I}_{\alpha,\tau} = \frac{1}{\sigma_\mu^2} \underset{N\to\infty}{\text{plim}}\, \frac{1}{N} \sum_{i=1}^{N} \mathbf{h}_i'$$

and if $N \to \infty$ or $N, T \to \infty$ such that $\frac{N}{T} \to \infty$

$$\text{plim}\, \frac{1}{\sqrt{N}} \mathcal{I}_{\alpha,\tau} = \mathbf{0}$$

If $N, T \to \infty$ simultaneously

$$\begin{aligned} \text{plim}\, \frac{1}{N} \mathcal{I}_{\alpha,\tau} &= \frac{1}{\sigma_u^2 \left(1-\rho_\lambda\right)^{-2} + \sigma_\mu^2} \underset{N\to\infty}{\text{plim}}\, \frac{1}{N} \sum_{i=1}^{N} \mathbf{h}_i' \\ \text{plim}\, \frac{1}{\sqrt{NT}} \mathcal{I}_{\pi,\tau} &= \frac{1}{\sigma_u^2 \left(1-\rho_\lambda\right)^{-2} + \sigma_\mu^2} \text{plim}\, \frac{1}{NT} \sum_{t=2}^{T-1} \mathbf{d}_t \sum_{i=1}^{N} \mathbf{h}_i' \end{aligned}$$

and otherwise for the last term

$$\text{plim}\, \frac{1}{\sqrt{NT}} \mathcal{I}_{\pi,\tau} = \mathbf{0}$$

Proof. These results can be proved with exactly the same methods as in lemma 6. In fact the same matrices are involved in the expressions and the proof is therefore omitted ■

Next we give the proofs of the theorems in the text.

Proof theorem 1. The method of proof is to examine the probability limit of the standardized log-likelihood. It is however not useful for dealing with the constant. In fact, the constant drops out of the analysis. The reason for adopting this method is that we can (in most cases) prove global consistency results for the other parameters which are not easily obtained otherwise. Asymptotic properties of the constant term are established separately at the end of the proof.

The negative of the log likelihood is up to an irrelevant term given by

$$\begin{aligned}\phi(\boldsymbol{\delta},\boldsymbol{\gamma}) &= \frac{N}{2}\ln\left|\mathbf{A}^*\right| + \frac{1}{2}\ln\left|\mathbf{I}_T + N\sigma_u^2\boldsymbol{\Psi}_\lambda\mathbf{A}^*\right| \\ &\quad + \frac{1}{2}\left(\mathbf{y}-\mathbf{Z}\boldsymbol{\delta}\right)'\boldsymbol{\Sigma}^{-1}\left(\mathbf{y}-\mathbf{Z}\boldsymbol{\delta}\right) \\ &= \frac{N}{2}\ln\left|\mathbf{A}^*\right| + \frac{1}{2}\ln\left|\mathbf{I}_T + N\sigma_u^2\boldsymbol{\Psi}_\lambda\mathbf{A}^*\right| \\ &\quad + \frac{1}{2}\left(\boldsymbol{\delta}_0-\boldsymbol{\delta}\right)'\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\mathbf{Z}\left(\boldsymbol{\delta}_0-\boldsymbol{\delta}\right) \\ &\quad + \frac{1}{2}\boldsymbol{\varepsilon}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} + \left(\boldsymbol{\delta}_0-\boldsymbol{\delta}\right)'\mathbf{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}\end{aligned} \tag{10}$$

By theorem 4.1.1 of Amemiya (1985) we need to verify that (i) the parameter space Θ is a compact subset of the Euclidean K-space, (ii) $\phi(\boldsymbol{\delta},\boldsymbol{\gamma})$ is continuous in $\theta\in\Theta$ for all $(\mathbf{y},\mathbf{X})$ and is a measurable function of $(\mathbf{y},\mathbf{X})$ for all $\theta\in\Theta$, (iii) $W^{-1}\phi(\boldsymbol{\delta},\boldsymbol{\gamma})$ converges to a nonstochastic function, say ϕ_0, in probability uniformly in $\theta\in\Theta$ as $W\to\infty$ and ϕ_0 is uniquely minimized at θ_0. Since (i) follows from assumption (b) and (ii) is trivial it remains to show (iii). This involves finding the limit of $W^{-1}\phi(\boldsymbol{\delta},\boldsymbol{\gamma})$ as $W\to\infty$ with $W=N$, $W=T$ and $W=NT$ respectively.

First we consider the uniform probability limit of (10) as $N,T\to\infty$. Note that

$$E\boldsymbol{\varepsilon}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} = \operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0$$

where $\boldsymbol{\Sigma}_0$ denotes $\boldsymbol{\Sigma}$ evaluated at θ_0. Hence using lemma 5, lemma 6 and 7 and assumption (c)

$$\begin{aligned}\operatorname*{plim}_{N,T\to\infty}\frac{1}{NT}\phi(\boldsymbol{\delta},\boldsymbol{\gamma}) &= -\lim_{T\to\infty}\frac{1}{2T}\ln\left|\mathbf{C}\right|^2 + \lim_{T\to\infty}\frac{1}{2T}\ln\left(d^2\sigma_\mu^2\left(1-\rho_v\right)^2+\sigma_e^2\right) \\ &\quad + \lim_{N,T\to\infty}\frac{N\left(T-1\right)}{2NT}\ln\sigma_e^2 \\ &\quad + \lim_{N,T\to\infty}\frac{1}{2NT}\ln\left|\mathbf{I}_T + N\sigma_u^2\boldsymbol{\Psi}_\lambda\mathbf{A}^*\right| \\ &\quad + \frac{1}{2}\left(\boldsymbol{\beta}_0-\boldsymbol{\beta}\right)'\mathbf{R}_{\mathbf{x}}\left(\boldsymbol{\beta}_0-\boldsymbol{\beta}\right) + \frac{\sigma_{e0}^2}{2\sigma_e^2}\left[\frac{\left(\rho_v^2+1\right)-2\rho_{v0}\rho_v}{1-\rho_{v0}^2}\right]\end{aligned}$$

with $\mathbf{R_x} = \text{plim}_{N,T\to\infty} \frac{1}{NT}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X}$, since $|\mathbf{C}| = O(1)$,

$$\lim_{N,T\to\infty} \frac{1}{2NT} \ln \left|\mathbf{I}_T + N\sigma_u^2 \mathbf{\Psi}_\lambda \mathbf{A}^*\right| = 0$$

and using lemma 4 we arrive at

$$\frac{1}{2}\ln\sigma_e^2 + \frac{1}{2}(\boldsymbol{\beta}_0 - \boldsymbol{\beta})' \mathbf{R_x} (\boldsymbol{\beta}_0 - \boldsymbol{\beta}) + \frac{\sigma_{e0}^2}{2\sigma_e^2}\left[\frac{(\rho_v^2+1) - 2\rho_{v0}\rho_v}{1-\rho_{v0}^2}\right] \tag{11}$$

and it is straightforward to verify that (11) is uniquely minimized at $\boldsymbol{\beta} = \boldsymbol{\beta}_0$, $\sigma_e^2 = \sigma_{e0}^2$ and $\mathbf{C} = \mathbf{C}_0$. Having established the consistency of maximum likelihood estimators $\widehat{\boldsymbol{\beta}}, \widehat{\sigma}_e^2, \widehat{\rho}_v$ as $N, T \to \infty$ we obtain the uniform probability limit of (10) as $N \to \infty$ with $T \geq 2$ a fix constant. For this purpose let $\boldsymbol{\zeta} = (\boldsymbol{\beta}', \boldsymbol{\tau}')'$

$$\begin{aligned}
\underset{N\to\infty}{\text{plim}} \frac{1}{NT}\phi(\boldsymbol{\delta},\boldsymbol{\gamma}) &= \lim_{N\to\infty}\frac{1}{2T}\ln|\mathbf{A}^*| + \lim_{N\to\infty}\frac{1}{2NT}\ln\left|\mathbf{I}_T + N\sigma_u^2\mathbf{\Psi}_\lambda\mathbf{A}^*\right| \\
&\quad + \underset{N\to\infty}{\text{plim}}\frac{1}{2NT}(\boldsymbol{\zeta}_0 - \boldsymbol{\zeta})'\mathbf{Z}_N'\mathbf{\Sigma}^{-1}\mathbf{Z}_N(\boldsymbol{\zeta}_0 - \boldsymbol{\zeta}) \\
&\quad + \lim_{N\to\infty}\frac{1}{2NT}E\boldsymbol{\varepsilon}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon} \\
&\quad + \underset{N\to\infty}{\text{plim}}\frac{1}{NT}(\boldsymbol{\zeta}_0 - \boldsymbol{\zeta})'\mathbf{Z}_N'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon} \\
&= -\frac{1}{2T}\ln\left(1-\rho_v^2\right) + \frac{1}{2T}\ln\left(d^2\sigma_\mu^2(1-\rho_v)^2 + \sigma_e^2\right) \\
&\quad + \left(\frac{1}{2} - \frac{1}{2T}\right)\ln\sigma_e^2 \\
&\quad + \frac{\sigma_{\mu0}^2 d^2(1-\rho_v)^2}{2T\left(d^2\sigma_\mu^2(1-\rho_v)^2 + \sigma_e^2\right)} + \frac{1}{2T}\sigma_{e0}^2\,\text{tr}\left(\mathbf{A}^*\mathbf{\Psi}_{v0}\right)
\end{aligned}$$

where we have used lemma 5 to evaluate $\lim_{N\to\infty}\frac{1}{2NT}E\boldsymbol{\varepsilon}'\mathbf{\Sigma}^{-1}\boldsymbol{\varepsilon}$, lemma 4, 6 and lemma 7 and that $\boldsymbol{\zeta}$ is uniquely identified with $\boldsymbol{\zeta} = \boldsymbol{\zeta}_0$. We then have

$$\begin{aligned}
\underset{N\to\infty}{\text{plim}} \frac{1}{NT}\phi(\boldsymbol{\delta},\boldsymbol{\gamma}) &= -\frac{1}{2T}\ln\left(1-\rho_v^2\right) + \frac{1}{2}\ln\sigma_e^2 \\
&\quad + \frac{1}{2}\sigma_{e0}^2\sigma_e^{-2}\frac{1}{T}\,\text{tr}\left(\mathbf{C}\mathbf{\Psi}_{v0}\mathbf{C}'\right) - \frac{1}{2T}\ln\sigma_e^2 \\
&\quad - \frac{1}{2}\sigma_{e0}^2\sigma_e^{-2}\frac{d^{-2}}{T}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\mathbf{\Psi}_{v0}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha} \\
&\quad + \frac{1}{2T}\ln\left(d^2\sigma_\mu^2(1-\rho_v)^2 + \sigma_e^2\right) \\
&\quad + \frac{1}{2T}\frac{\sigma_{\mu0}^2 d^2(1-\rho_v)^2 + \sigma_{e0}^2 d^{-2}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\mathbf{\Psi}_{v0}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha}}{\left(d^2\sigma_\mu^2(1-\rho_v)^2 + \sigma_e^2\right)}
\end{aligned} \tag{12}$$

Evaluating $\mathrm{tr}\left(\mathbf{C}\boldsymbol{\Psi}_{v0}\mathbf{C}'\right)$ and $\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{C}\boldsymbol{\Psi}_{v0}\mathbf{C}'\boldsymbol{\iota}_T^{\alpha}$ as in lemma 3 we can show that (12) is uniquely minimized at $\rho_v = \rho_{v0}, \sigma_e^2 = \sigma_{e0}^2$ and $\sigma_\mu^2 = \sigma_{\mu 0}^2$. This establishes the consistency of $\widehat{\boldsymbol{\beta}}, \widehat{\sigma}_e^2, \widehat{\rho}_v$ as $N \to \infty$ as well as the consistency of $\widehat{\sigma}_\mu^2, \widehat{\boldsymbol{\tau}}$ as $N \to \infty$ or $N, T \to \infty$.

Consider next the uniform probability limit of (10) as $N \geq 2$ is fix and $T \to \infty$. Noting that lemma 6 and 7 and assumption (c) ensures that $\boldsymbol{\psi} = (\boldsymbol{\beta}', \boldsymbol{\pi}')'$ is uniquely identified we have

$$\begin{aligned}\operatorname*{plim}_{T\to\infty} \frac{1}{NT}\phi(\boldsymbol{\delta},\boldsymbol{\gamma}) &= \lim_{T\to\infty}\frac{1}{2T}\ln\left|\mathbf{A}^*\right| + \lim_{T\to\infty}\frac{1}{2NT}\ln\left|\mathbf{I}_T + N\sigma_u^2\boldsymbol{\Psi}_\lambda\mathbf{A}^*\right| \\ &\quad + \lim_{T\to\infty}\frac{1}{2NT}\,\mathrm{tr}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0\end{aligned}$$

using lemma 5 and after some matrix manipulation we arrive at

$$\begin{aligned}\operatorname*{plim}_{T\to\infty} \frac{1}{NT}\phi(\boldsymbol{\delta},\boldsymbol{\gamma}) &= \left(\frac{1}{2}-\frac{1}{2N}\right)\ln\sigma_e^2 \\ &\quad + \left(\frac{1}{2}-\frac{1}{2N}\right)\frac{\sigma_{e0}^2}{\sigma_e^2}\left[\frac{(\rho_v^2+1)-2\rho_{v0}\rho_v}{1-\rho_{v0}^2}\right] \\ &\quad + \frac{1}{2N}\lim_{T\to\infty}\frac{1}{T}\ln\left|N\sigma_u^2\boldsymbol{\Psi}_\lambda + \sigma_e^2\boldsymbol{\Psi}_v\right| \\ &\quad + \frac{1}{2N}\lim_{T\to\infty}\frac{1}{T}\,\mathrm{tr}\left(\mathbf{P}_0\mathbf{P}^{-1}\right)\end{aligned} \tag{13}$$

where $\mathbf{P}_0$ and $\mathbf{P}$ are given in lemma 5. The first and second row of (13) are uniquely minimized at $\sigma_e^2 = \sigma_{e0}^2$, $\rho_v = \rho_{v0}$. However we cannot evaluate the last two rows analytically which complicates showing uniqueness globally[6]. We can prove the existence of a consistent root though (cf. Amemiya (1985, theorem 4.1.2)). Applying matrix differentiation to (13) using standard results for interchanging the limit and the derivative e.g. Rudin (1976, p 152) it is straightforward to show that the true parameters are a solution to the first order condition. Of course then we also need to verify that the second derivative matrix is positive-definite when evaluated at the true parameters. But this is straightforward to do as well. This proves the global consistency

[6] In case of $\rho_v = \rho_\lambda = 0$ (13) reduces to

$$\begin{aligned}&\left(\frac{1}{2}-\frac{1}{2N}\right)\ln\sigma_e^2 + \left(\frac{1}{2}-\frac{1}{2N}\right)\frac{\sigma_{e0}^2}{\sigma_e^2} \\ &+\frac{1}{2N}\ln\left(N\sigma_u^2+\sigma_e^2\right) + \frac{1}{2N}\frac{N\sigma_{u0}^2+\sigma_{e0}^2}{N\sigma_u^2+\sigma_e^2}\end{aligned}$$

which is globally minimized at the true parameters if $N \geq 2$.

of $\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\pi}}$ as $T \to \infty$ (and also the global consistency of $\widehat{\boldsymbol{\pi}}$ as $N, T \to \infty$) and the existence of a local consistent root for $\widehat{\sigma}_e^2, \widehat{\rho}_v, \widehat{\sigma}_u^2, \widehat{\rho}_\lambda$ as $T \to \infty$. Since the information matrix is positive definite over the full parameter space when $N, T \to \infty$ (as shown in theorem 2) this also proves the global consistency of $\widehat{\sigma}_u^2, \widehat{\rho}_\lambda$ as $N, T \to \infty$.

Finally we obtain results for the constant term. To obtain a local consistency result for $\widehat{\alpha}$ as $N, T \to \infty$ it suffices to consider lemma 6. In fact $\widehat{\alpha}$ can be shown to be globally consistent as $N, T \to \infty$ by the results in lemma 6 and lemma 7 and the fact that the information matrix is positive definite over the full parameter space for the remaining parameters. As a special case of lemma 6 we obtain the inconsistency of $\widehat{\alpha}$ as only $N \to \infty$ or $T \to \infty$ ■

Proof theorem 2. We first derive the results when $N, T \to \infty$ and hence the full parameter vector is consistently estimated. For the purpose of establishing asymptotic normality of $\widehat{\theta}$ it is useful to structure $\boldsymbol{\delta}$ as $\boldsymbol{\delta} = (\alpha, \boldsymbol{\pi}', \boldsymbol{\tau}', \boldsymbol{\beta}')'$ and we will do so below. By the mean value theorem for random functions Gourieroux and Monfort (1995, p 400)

$$\frac{\partial l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial \theta}|_{\widehat{\theta}} = \frac{\partial l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial \theta}|_{\theta_0} + \frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial \theta \partial \theta'}|_{\overline{\theta}} \left(\widehat{\theta} - \theta_0\right) \tag{14}$$

where $\overline{\theta}$ belongs to the segment $\left(\widehat{\theta}, \theta_0\right)$ with probability 1. Define $\mathbf{F}_{NT}$ as a diagonal matrix with

$$\operatorname{diag} \mathbf{F}_{NT} = \left\{\min\left(\sqrt{N}, \sqrt{T}\right), \mathbf{F}_\pi, \mathbf{F}_\tau, \mathbf{F}_\beta, \sqrt{N}, \sqrt{NT}, \sqrt{NT}, \sqrt{T}, \sqrt{T}\right\}$$

where $\mathbf{F}_\beta$ is a vector containing k_1 $\sqrt{NT}$ and $\mathbf{F}_\pi, \mathbf{F}_\tau$ are vectors containing k_1 $\sqrt{T}$ and k_2 $\sqrt{N}$ respectively. We can then write

$$\mathbf{F}_{NT}\left(\widehat{\theta} - \theta_0\right) = -\left[\mathbf{F}_{NT}^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial \theta \partial \theta'}|_{\overline{\theta}}\right)\mathbf{F}_{NT}^{-1}\right]^{-1}\left[\mathbf{F}_{NT}^{-1}\left(\frac{\partial l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial \theta}|_{\theta_0}\right)\right] \tag{15}$$

From theorem 4.1.3 of Amemiya (1985) we need to show that (in addition to local consistency Amemiya (1985, theorem 4.1.2)) (i) $l(\boldsymbol{\delta}, \boldsymbol{\gamma}) \in C^2$ in a convex neighborhood of θ_0, (ii) $\left[\mathbf{F}_{NT}^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial \theta \partial \theta'}|_{\overline{\theta}}\right)\mathbf{F}_{NT}^{-1}\right]$ converges to a finite non-singular matrix

$$\mathbf{V}^{-1}(\theta_0) = -\lim_{N,T\to\infty} E\left[\mathbf{F}_{NT}^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial \theta \partial \theta'}|_{\theta_0}\right)\mathbf{F}_{NT}^{-1}\right]$$

in probability for any sequence $\overline{\theta}$ such that $\operatorname{plim} \overline{\theta} = \theta_0$ and (iii)

$$\mathbf{F}_{NT}^{-1}\left(\frac{\partial l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial \theta}|_{\theta_0}\right) \xrightarrow{d} N\left(\mathbf{0}, \mathbf{V}_1^{-1}(\theta_0)\right)$$

where

$$\mathbf{V}_1^{-1}(\theta_0) = \lim_{N,T\to\infty} E\left[\mathbf{F}_{NT}^{-1}\left(\frac{\partial l(\boldsymbol{\delta},\boldsymbol{\gamma})}{\partial\theta}|_{\theta_0}\right)\left(\frac{\partial l(\boldsymbol{\delta},\boldsymbol{\gamma})}{\partial\theta}|_{\theta_0}\right)'\mathbf{F}_{NT}^{-1}\right]$$

a finite non-singular matrix. Note that (i) is trivially satisfied and by assumption (a) (ii) follows if the convergence is uniform. Further note that

$$E\left[\left(\frac{\partial l(\boldsymbol{\delta},\boldsymbol{\gamma})}{\partial\theta}|_{\theta_0}\right)\right] = 0$$

is straightforward to verify from appendix A.1, and

$$\mathbf{V}_1(\theta_0) = \mathbf{V}(\theta_0)$$

follows from the information matrix equality. To show (ii) we take uniform limits of the appropriately scaled elements of the information matrix obtained from appendix A.2. The limits for the variance parameters are straightforward to derive using lemma 4 and repeatedly using elementary results on inverses involving sums. For the elements $\mathcal{I}_{\sigma_\mu^2,\gamma_j}$ we have

$$\begin{aligned}
\lim_{N,T\to\infty}\frac{1}{N}\mathcal{I}_{\sigma_\mu^2,\sigma_\mu^2} &= \lim_{N,T\to\infty}\frac{1}{2}\left[\frac{d^4}{\sigma_\alpha^4}(1-\rho_v)^4\right] \\
&= \frac{1}{2}\lim_{N,T\to\infty}\frac{\left(\alpha^2+(T-1)\right)^2(1-\rho_v)^4}{\left(\left(\alpha^2+(T-1)\right)\sigma_\mu^2(1-\rho_v)^2+\sigma_e^2\right)^2} \\
&= \frac{1}{2\sigma_\mu^4}
\end{aligned}$$

$$\lim_{N,T\to\infty}\frac{1}{N\sqrt{T}}\mathcal{I}_{\sigma_\mu^2,\sigma_e^2} = 0$$

$$\lim_{N,T\to\infty}\frac{1}{\sqrt{NT}}\mathcal{I}_{\sigma_\mu^2,\sigma_u^2} = 0$$

$$\lim_{N,T\to\infty}\frac{1}{\sqrt{NT}}\mathcal{I}_{\sigma_\mu^2,\rho_\lambda} = 0$$

$$\lim_{N,T\to\infty}\frac{1}{N\sqrt{T}}\mathcal{I}_{\sigma_\mu^2,\rho_v} = 0$$

and for the elements $\mathcal{I}_{\sigma_e^2,\gamma_j}$

$$\begin{aligned}
\lim_{N,T\to\infty}\frac{1}{NT}\mathcal{I}_{\sigma_e^2,\sigma_e^2} &= \lim_{N,T\to\infty}\frac{1}{2T}\operatorname{tr}\left(\mathbf{A}^*\boldsymbol{\Psi}_v\right)^2 \\
&= \frac{1}{2\sigma_e^4}
\end{aligned}$$

$$\lim_{N,T\to\infty}\frac{1}{\sqrt{NT}}\mathcal{I}_{\sigma_e^2,\sigma_u^2}=0$$

$$\lim_{N,T\to\infty}\frac{1}{\sqrt{NT}}\mathcal{I}_{\sigma_e^2,\rho_\lambda}=0$$

$$\lim_{N,T\to\infty}\frac{1}{NT}\mathcal{I}_{\sigma_e^2,\rho_v}=\lim_{N,T\to\infty}\frac{1}{2\sigma_e^2 T}\operatorname{tr}\left(\mathbf{\Psi}_v^{-1}\mathbf{L}_v\right)=0$$

Finally, for the elements involving $\sigma_u^2,\rho_\lambda,\rho_v$

$$\lim_{N,T\to\infty}\frac{1}{T}\mathcal{I}_{\sigma_u^2,\sigma_u^2}=\frac{1}{2\sigma_u^4}$$

$$\begin{aligned}
\lim_{N,T\to\infty}\frac{1}{T}\mathcal{I}_{\sigma_u^2,\rho_\lambda} &= \lim_{N,T\to\infty}\frac{1}{2\sigma_u^2 T}\operatorname{tr}\left(\mathbf{\Psi}_\lambda^{-1}\mathbf{L}_\lambda\right)\\
&= \frac{\rho_\lambda}{\sigma_u^2\left(1-\rho_\lambda^2\right)}+\lim_{N,T\to\infty}\frac{1}{2\sigma_u^2 T\left(1-\rho_\lambda^2\right)}\operatorname{tr}\left(\mathbf{\Psi}_\lambda^{-1}\mathbf{D}\right)\\
&= \frac{\rho_\lambda}{\sigma_u^2\left(1-\rho_\lambda^2\right)}+\lim_{N,T\to\infty}\frac{\rho_\lambda}{\sigma_u^2 T\left(1-\rho_\lambda^2\right)}(1-T)=0
\end{aligned}$$

$$\lim_{N,T\to\infty}\frac{1}{\sqrt{NT}}\mathcal{I}_{\sigma_u^2,\rho_v}=0$$

$$\begin{aligned}
\lim_{N,T\to\infty}\frac{1}{T}\mathcal{I}_{\rho_\lambda,\rho_\lambda} &= \lim_{N,T\to\infty}\frac{1}{2T}\operatorname{tr}\left(\mathbf{\Psi}_\lambda^{-1}\mathbf{L}_\lambda\right)^2\\
&= \lim_{N,T\to\infty}\frac{1}{2T}\operatorname{tr}\left(\mathbf{\Psi}_\lambda^{-1}\left(\frac{2\rho_\lambda}{\left(1-\rho_\lambda^2\right)}\mathbf{\Psi}_\lambda+\frac{1}{\left(1-\rho_\lambda^2\right)}\mathbf{D}\right)\right)^2\\
&= -\frac{2\rho_\lambda^2}{\left(1-\rho_\lambda^2\right)^2}+\lim_{N,T\to\infty}\frac{1}{2T\left(1-\rho_\lambda^2\right)^2}\operatorname{tr}\left(\mathbf{\Psi}_\lambda^{-1}\mathbf{D}\right)^2\\
&= \frac{1}{\left(1-\rho_\lambda^2\right)}
\end{aligned}$$

$$\lim_{N,T\to\infty}\frac{1}{\sqrt{NT}}\mathcal{I}_{\rho_\lambda,\rho_v}=0$$

$$\begin{aligned}
\lim_{N,T\to\infty}\frac{1}{NT}\mathcal{I}_{\rho_v,\rho_v} &= \lim_{N,T\to\infty}\frac{\sigma_e^4}{2T}\operatorname{tr}\left(\mathbf{A}^*\mathbf{L}_v\right)^2=\lim_{N,T\to\infty}\frac{1}{2T}\operatorname{tr}\left(\mathbf{C}'\mathbf{C}\mathbf{L}_v\right)^2\\
&= \lim_{N,T\to\infty}\frac{1}{2T}\operatorname{tr}\left(\mathbf{\Psi}_v^{-1}\mathbf{L}_v\right)^2=\frac{1}{\left(1-\rho_v^2\right)}
\end{aligned}$$

where we have used that $\mathbf{L}_\lambda=\frac{\partial\mathbf{\Psi}_\lambda}{\partial\rho_\lambda}=\frac{2\rho_\lambda}{\left(1-\rho_\lambda^2\right)}\mathbf{\Psi}_\lambda+\frac{1}{\left(1-\rho_\lambda^2\right)}\mathbf{D}$ with $\mathbf{D}$ a band matrix with zeros on the main diagonal and $i\rho_\lambda^{i-1}$ on the i:th subdiagonal and tr $\left(\boldsymbol{\Psi}_\lambda^{-1}\mathbf{D}\right) = 2\rho_\lambda (1-T)$, tr $\left(\boldsymbol{\Psi}_\lambda^{-1}\mathbf{D}\right)^2 = 2\rho_\lambda^2 (T-1) + 2(T-1)$. Hence we arrive at (using assumption (c), lemma 6 and lemma 7)

$$\mathbf{V}^{-1}(\theta_0) = \begin{bmatrix} \boldsymbol{\phi} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_X} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}^{-1}(\theta_0)_\gamma \end{bmatrix}$$

where $\mathbf{R_X} = \text{plim}_{N,T\to\infty} \frac{1}{NT}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X}$, and with

$$\boldsymbol{\phi} = \underset{T\to\infty}{\text{plim}} \begin{bmatrix} \frac{(1-\rho_{\lambda 0})^2}{\sigma_{u0}^2} & \frac{(1-\rho_\lambda)^2}{T\sigma_{u0}^2}\sum_{t=2}^{T-1}\mathbf{d}_t' & \mathbf{0} \\ & \frac{1}{T\sigma_{u0}^2}\sum_{t=2}^{T}\mathbf{d}_t^\lambda\mathbf{d}_t^{\lambda\prime} & \mathbf{0} \\ & & \frac{1}{\sigma_{\mu 0}^2}\text{plim}_{N\to\infty}\frac{1}{N}\mathbf{h}'\overline{\mathbf{E}}_N\mathbf{h} \end{bmatrix}$$

if $\frac{N}{T} \to \infty$ where $\mathbf{d}_t^\lambda = (\mathbf{d}_t - \rho_\lambda \mathbf{d}_{t-1})$, $\overline{\mathbf{E}}_N = \mathbf{I}_N - \overline{\mathbf{J}}_N, \overline{\mathbf{J}}_N = \frac{1}{N}\boldsymbol{\iota}_N\boldsymbol{\iota}_N'$

$$\boldsymbol{\phi} = \underset{N\to\infty}{\text{plim}} \begin{bmatrix} \frac{1}{\sigma_{\mu 0}^2} & \mathbf{0} & \frac{1}{N\sigma_{\mu 0}^2}\sum_{i=1}^{N}\mathbf{h}_i' \\ & \text{plim}_{T\to\infty}\frac{1}{T\sigma_{u0}^2}\mathbf{S}_\lambda & \mathbf{0} \\ & & \frac{1}{N\sigma_{\mu 0}^2}\mathbf{h}'\mathbf{h} \end{bmatrix}$$

if $\frac{T}{N} \to \infty$ where $\mathbf{S}_\lambda = \left(\sum_{t=2}^{T}\mathbf{d}_t^\lambda\mathbf{d}_t^{\lambda\prime} - \frac{(1-\rho_{\lambda 0})^2}{T}\sum_{t=2}^{T-1}\sum_{r=2}^{T-1}\mathbf{d}_t\mathbf{d}_r'\right)$, and

$$\boldsymbol{\phi} = \underset{z\to\infty}{\text{plim}} \begin{bmatrix} \omega & \omega\frac{1}{T}\sum_{t=2}^{T-1}\mathbf{d}_t' & \omega\frac{1}{N}\sum_{i=1}^{N}\mathbf{h}_i' \\ & \frac{1}{T\sigma_{u0}^2}\mathbf{S}_\lambda + \omega\frac{1}{T^2}\sum_{t=2}^{T-1}\sum_{r=2}^{T-1}\mathbf{d}_t\mathbf{d}_r' & \frac{\omega}{NT}\sum_{t=2}^{T-1}\mathbf{d}_t\sum_{i=1}^{N}\mathbf{h}_i' \\ & & \frac{1}{N\sigma_{\mu 0}^2}\mathbf{h}'\overline{\mathbf{E}}_N\mathbf{h} + \omega\frac{1}{N}\mathbf{h}'\overline{\mathbf{J}}_N\mathbf{h} \end{bmatrix}$$

if $N, T \to \infty$ simultaneously, where $\omega = \frac{1}{\sigma_{\mu 0}^2 + (1-\rho_{\lambda 0})^{-2}\sigma_{u0}^2}$, $z = N$ or T and $\mathbf{V}^{-1}(\boldsymbol{\theta}_0)_\gamma$ is a diagonal matrix with

$$\text{diag}\,\mathbf{V}^{-1}(\theta_0)_\gamma = \left\{\frac{1}{2\sigma_{\mu 0}^4}, \frac{1}{2\sigma_{e0}^4}, \frac{1}{(1-\rho_{v0}^2)}, \frac{1}{2\sigma_{u0}^4}, \frac{1}{(1-\rho_{\lambda 0}^2)}\right\}$$

To show (iii) note that the elements of the score for $\boldsymbol{\delta}$ is a linear combination of the normal $\boldsymbol{\varepsilon}$ and the score for the variance parameters, $\boldsymbol{\gamma}$ are linear combinations of quadratic forms in normal variates i.e. they can be written as

$$b + \boldsymbol{\varepsilon}'\mathbf{P}\boldsymbol{\varepsilon}$$

for suitable choice of b and symmetric matrix $\mathbf{P}$. We then apply the following lemma adapted from Amemiya (1971) to the quadratic forms in appendix A.1

Lemma. *Let an n-component vector random variable $u \sim N(\mathbf{0}, \mathbf{\Lambda})$, $\mathbf{G}$ be a non-negative definite symmetric matrix with rank $r \leq n$. Then $\mathbf{u}'\mathbf{G}\mathbf{u}$ is distributed as $\sum_{i=1}^{r} \varphi_i \chi_i^2(1)$, where the φ's are r non-zero characteristic roots of $\mathbf{\Lambda G}$ and each $\chi_i^2(1)$ is an independent chi-square. If $\mathbf{H}$ is another non-negative definite symmetric matrix,* $\operatorname{cov}(\mathbf{u}'\mathbf{G}\mathbf{u}, \mathbf{u}'\mathbf{H}\mathbf{u}) = 2\operatorname{tr}(\mathbf{G\Lambda H\Lambda})$.

Asymptotic normality of the appropriately normalized score vector can then be shown by establishing sequential weak convergence results in case $\frac{N}{T} \to \infty$ or $\frac{T}{N} \to \infty$ (see Phillips and Moon (1999, section 3.3)) and in case $N, T \to \infty$ simultaneously a multivariate CLT for triangular arrays may be applied.

To establish the results as only N or $T \to \infty$ we apply the expansion (15) to the consistent subvectors $\theta^i = \left(\boldsymbol{\beta}, \boldsymbol{\tau}, \boldsymbol{\gamma}^{(i)}\right), \boldsymbol{\gamma}^{(i)} = \left(\sigma_\mu^2, \sigma_e^2, \rho_v\right)$ and $\theta^t = \left(\boldsymbol{\beta}, \boldsymbol{\pi}, \boldsymbol{\gamma}^{(t)}\right), \boldsymbol{\gamma}^{(t)} = \left(\sigma_e^2, \rho_v, \sigma_u^2, \rho_\lambda\right)$ as N and $T \to \infty$ respectively. This gives

$$\mathbf{F}_N\left(\widehat{\theta}^i - \theta_0^i\right) = -\left[\mathbf{F}_N^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial\theta^i \partial\theta^{i\prime}}|_{\overline{\theta}^i}\right)\mathbf{F}_N^{-1}\right]^{-1}\left[\mathbf{F}_N^{-1}\left(\frac{\partial l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial\theta^i}|_{\theta_0^i}\right)\right]$$

and

$$\mathbf{F}_T\left(\widehat{\theta}^t - \theta_0^t\right) = -\left[\mathbf{F}_T^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial\theta^t \partial\theta^{t\prime}}|_{\overline{\theta}^t}\right)\mathbf{F}_T^{-1}\right]^{-1}\left[\mathbf{F}_T^{-1}\left(\frac{\partial l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial\theta^t}|_{\theta_0^t}\right)\right]$$

where $\mathbf{F}_N, \mathbf{F}_T$ are diagonal matrices with

$$\begin{aligned}
\operatorname{diag}\mathbf{F}_N &= \left\{\mathbf{F}_\beta, \mathbf{F}_\tau, \sqrt{N}, \sqrt{NT}, \sqrt{NT}\right\} \\
\operatorname{diag}\mathbf{F}_T &= \left\{\mathbf{F}_\beta, \mathbf{F}_\pi, \sqrt{NT}, \sqrt{NT}, \sqrt{T}, \sqrt{T}\right\}
\end{aligned}$$

To show (ii) for these cases we need to examine the convergence of the information matrices as N and $T \to \infty$ respectively. As $N \to \infty$ we find (using assumption (c), lemma 6, lemma 7 and straightforward computations)

$$\lim_{N\to\infty} E\left[\mathbf{F}_N^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta}, \boldsymbol{\gamma})}{\partial\theta^i \partial\theta^{i\prime}}|_{\theta_0^i}\right)\mathbf{F}_N^{-1}\right] = \begin{bmatrix} \mathbf{R}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_N^{-1}\left(\theta_0^i\right)_{\gamma^{(i)}} \end{bmatrix}$$

where

$$\begin{aligned}
&\mathbf{V}_N^{-1}\left(\theta_0^i\right)_{\gamma^{(i)}} \\
&= \frac{1}{2}\begin{bmatrix} \left(\frac{\sigma_\varpi^2}{\sigma_\alpha^2\sigma_\mu^2}\right)^2 & \frac{(1-\rho_v)^2}{\sigma_\alpha^2\sqrt{T}}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{\Psi}_v\boldsymbol{\iota}_T^\alpha & \frac{\sigma_e^2(1-\rho_v)^2}{\sigma_\alpha^2\sqrt{T}}\boldsymbol{\iota}_T^{\alpha\prime}\mathbf{A}^*\mathbf{L}_v\boldsymbol{\iota}_T^\alpha \\ & \frac{1}{T}\left(\sigma_\alpha^{-4} + (T-1)\,\sigma_e^{-4}\right) & \frac{\sigma_e^2}{T}\operatorname{tr}\left(\mathbf{A}^*\mathbf{\Psi}_v\mathbf{A}^*\mathbf{L}_v\right) \\ & & \frac{\sigma_e^4}{T}\operatorname{tr}\left(\mathbf{A}^*\mathbf{L}_v\right)^2 \end{bmatrix}
\end{aligned}$$

where $\sigma^2_{\varpi} = (\sigma^2_{\alpha} - \sigma^2_e)$ and $\mathbf{V}_N^{-1}\left(\theta_0^i\right)_{\gamma(i)}$ is positive-definite by theorem 1 and standard results of multivariate calculus.

As $T \rightarrow \infty$ we have (using assumption (c), lemma 6, lemma 7 and straightforward computations again)

$$\lim_{T\rightarrow\infty} E\left[\mathbf{F}_T^{-1}\left(\frac{\partial^2 l(\boldsymbol{\delta},\boldsymbol{\gamma})}{\partial\theta^t\partial\theta^{t\prime}}|_{\theta_0^t}\right)\mathbf{F}_T^{-1}\right] = \begin{bmatrix} \mathbf{R}_T & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_T^{-1}\left(\theta_0^t\right)_{\gamma(t)} \end{bmatrix}$$

with

$$\mathbf{V}_T^{-1}\left(\theta_0^t\right)_{\gamma(t)} = \lim_{T\rightarrow\infty}\frac{1}{2}\begin{bmatrix} \frac{1}{\sigma_e^4 T}\mathbf{V}^N_{\mathbf{\Psi}_v,\mathbf{\Psi}_v} & \frac{1}{\sigma_e^2 T}\mathbf{V}^N_{\mathbf{L}_v,\mathbf{\Psi}_v} & \frac{\sqrt{N}}{\sigma_e^4 T}\mathbf{V}_{\mathbf{\Psi}_v,\mathbf{\Psi}_\lambda} & \frac{\sigma_u^2\sqrt{N}}{\sigma_e^4 T}\mathbf{V}_{\mathbf{\Psi}_v,\mathbf{L}_\lambda} \\ & \frac{1}{T}\mathbf{V}^N_{\mathbf{L}_v,\mathbf{L}_v} & \frac{\sqrt{N}}{\sigma_e^2 T}\mathbf{V}_{\mathbf{L}_v,\mathbf{\Psi}_\lambda} & \frac{\sigma_u^2\sqrt{N}}{\sigma_e^2 T}\mathbf{V}_{\mathbf{L}_v,\mathbf{L}_\lambda} \\ & & \frac{N^2}{T\sigma_e^4}\mathbf{V}_{\mathbf{\Psi}_\lambda,\mathbf{\Psi}_\lambda} & \frac{\sigma_u^2 N^2}{\sigma_e^4 T}\mathbf{V}_{\mathbf{L}_\lambda,\mathbf{\Psi}_\lambda} \\ & & & \frac{\sigma_u^4 N^2}{T\sigma_e^4}\mathbf{V}_{\mathbf{L}_\lambda,\mathbf{L}_\lambda} \end{bmatrix}$$

where

$$\begin{aligned} \mathbf{V}_{\mathbf{F},\mathbf{P}} &= \operatorname{tr}\left(\left(\mathbf{\Psi}_v^{-1}\mathbf{F}\mathbf{\Psi}_v^{-1}\mathbf{P}\right)(\mathbf{I}_T - 2\mathbf{M})\right) + \operatorname{tr}\left(\mathbf{\Psi}_v^{-1}\mathbf{F}\mathbf{M}\mathbf{\Psi}_v^{-1}\mathbf{P}\mathbf{M}\right) \\ \mathbf{V}^N_{\mathbf{F},\mathbf{P}} &= \operatorname{tr}\left(\left(\mathbf{\Psi}_v^{-1}\mathbf{F}\mathbf{\Psi}_v^{-1}\mathbf{P}\right)\left(\mathbf{I}_T - \frac{2}{N}\mathbf{M}\right)\right) + \frac{1}{N}\operatorname{tr}\left(\mathbf{\Psi}_v^{-1}\mathbf{F}\mathbf{M}\mathbf{\Psi}_v^{-1}\mathbf{P}\mathbf{M}\right) \end{aligned}$$

and $\mathbf{M} = \left(\mathbf{I}_T + \frac{\sigma_e^2}{N\sigma_u^2}\mathbf{\Psi}_\lambda\mathbf{\Psi}_v\mathbf{\Psi}_\lambda^{-2}\right)^{-1}$. The positive-definiteness of $\mathbf{V}_T^{-1}\left(\theta_0^t\right)_{\gamma(t)}$ then follows from the results in theorem 1.

These results show that the information elements of the subsets of consistent variance parameters do not depend on the inconsistent nuisance parameters as $N \rightarrow \infty$ and $T \rightarrow \infty$ respectively. To show this for the information elements of the subsets of consistent regression parameters as well we write, as in lemma 6,

$$\begin{aligned} & \frac{1}{NT}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X} \\ = \; & \frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{A}^*\mathbf{X}_i - \frac{1}{N^2T}\sum_{j=1}^N\sum_{i=1}^N \mathbf{X}_j'\mathbf{A}^*\mathbf{X}_i + \frac{1}{N^2T}\sum_{j=1}^N\sum_{i=1}^N \mathbf{X}_j'\mathbf{L}^{-1}\mathbf{X}_i \end{aligned}$$

where $\mathbf{L} = \left(N\sigma_u^2\mathbf{\Psi}_\lambda + (\mathbf{A}^*)^{-1}\right)^{-1} = O\left(N^{-1}\right)$. Hence

$$\begin{aligned} & \operatorname*{plim}_{N\rightarrow\infty}\frac{1}{NT}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X} \\ = \; & \operatorname*{plim}_{N\rightarrow\infty}\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{A}^*\mathbf{X}_i - \operatorname*{plim}_{N\rightarrow\infty}\frac{1}{N^2T}\sum_{j=1}^N\sum_{i=1}^N \mathbf{X}_j'\mathbf{A}^*\mathbf{X}_i = \mathbf{R}_1\left(\theta_0^i\right) \end{aligned}$$

and as $T \to \infty$ we find

$$\operatorname*{plim}_{T\to\infty} \frac{1}{NT}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X} = \mathbf{R}_2\left(\theta_0^t\right)$$

since $\mathbf{A}^* = \mathbf{C}'\left(\sigma_\alpha^{-2}\overline{\mathbf{J}}_T^\alpha + \sigma_e^{-2}\mathbf{E}_T^\alpha\right)\mathbf{C}$ where $\sigma_\alpha^2 = O(T)$. Similarly one can show that cross-elements as well as information elements of time-invariant explanatory variables and individual-invariant random variables do not depend on nuisance parameters as N and $T \to \infty$ respectively. Finally, asymptotic normality of the limiting score vectors (suitably normalized of course) follows from applying a suitable multivariate CLT ∎

Proof theorem 3. The negative of the log-likelihood is (apart from a constant term) given by

$$\begin{aligned}\phi(\boldsymbol{\delta},\boldsymbol{\gamma}) &= \frac{N}{2}\ln|\mathbf{A}^*| + \frac{1}{2}(\boldsymbol{\delta}_0-\boldsymbol{\delta})'\mathbf{Z}'(\mathbf{I}_N\otimes\mathbf{A}^*)\mathbf{Z}(\boldsymbol{\delta}_0-\boldsymbol{\delta}) \\ &\quad +\frac{1}{2}\boldsymbol{\varepsilon}'(\mathbf{I}_N\otimes\mathbf{A}^*)\boldsymbol{\varepsilon}+(\boldsymbol{\delta}_0-\boldsymbol{\delta})'\mathbf{Z}'(\mathbf{I}_N\otimes\mathbf{A}^*)\boldsymbol{\varepsilon}\end{aligned}$$

Since the parameters ρ_λ, ρ_v play no role in what follows we assume $\rho_\lambda = \rho_v = 0$. To prove (i) note that

$$\operatorname{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_0\right) = \frac{N\left(T\sigma_{\mu 0}^2+\sigma_{e0}^2\right)}{\left(T\sigma_\mu^2+\sigma_e^2\right)} + \frac{N(T-1)\sigma_{e0}^2}{\sigma_e^2} + \frac{N\sigma_{u0}^2}{\left(T\sigma_\mu^2+\sigma_e^2\right)} + \frac{N(T-1)\sigma_{u0}^2}{\sigma_e^2}$$

which contradicts a consistent root of σ_μ^2, σ_e^2 as $N \to \infty$ and a consistent root of σ_e^2 as $T \to \infty$ or $N, T \to \infty$ (and hence also of ρ_v as $N \to \infty, T \to \infty$ or $N, T \to \infty$). To show (ii) we need to investigate the behavior of $\mathbf{Z}'(\mathbf{I}_N\otimes\mathbf{A}^*)\mathbf{Z}$, $\mathbf{Z}'(\mathbf{I}_N\otimes\mathbf{A}^*)\boldsymbol{\varepsilon}$ which are explicitly written as

$$\begin{aligned}&\mathbf{Z}'(\mathbf{I}_N\otimes\mathbf{A}^*)\mathbf{Z} \\ = &\begin{bmatrix} \frac{NT}{\sigma_1^2} & \frac{1}{\sigma_1^2}\sum_{i=1}^N \boldsymbol{\iota}_T'\mathbf{X}_i & \frac{N}{\sigma_1^2}\boldsymbol{\iota}_T'\mathbf{d} & \frac{T}{\sigma_1^2}\sum_{i=1}^N h_i \\ & \sum_{i=1}^N \mathbf{X}_i'\mathbf{A}^*\mathbf{X}_i & \sum_{i=1}^N \mathbf{X}_i'\mathbf{A}^*\mathbf{d} & \frac{1}{\sigma_1^2}\sum_{i=1}^N \mathbf{X}_i'\boldsymbol{\iota}_T h_i \\ & & N\mathbf{d}'\mathbf{A}^*\mathbf{d} & \frac{1}{\sigma_1^2}\sum_{i=1}^N \mathbf{d}'\boldsymbol{\iota}_T h_i \\ & & & \frac{T}{\sigma_1^2}\sum_{i=1}^N h_i'h_i \end{bmatrix}\end{aligned}$$

and

$$\begin{aligned}&\mathbf{Z}'(\mathbf{I}_N\otimes\mathbf{A}^*)\boldsymbol{\varepsilon} \\ = &\begin{bmatrix} \frac{T}{\sigma_1^2}\sum_{i=1}^N \mu_i + \frac{N}{\sigma_1^2}\sum_{t=1}^T \lambda_t + \frac{1}{\sigma_1^2}\sum_{t=1}^T\sum_{i=1}^N v_{it} \\ \frac{1}{\sigma_1^2}\sum_{t=1}^T\sum_{i=1}^N X_{it}\mu_i + \sum_{i=1}^N \mathbf{X}_i'\mathbf{A}^*\boldsymbol{\lambda} + \sum_{i=1}^N \mathbf{X}_i'\mathbf{A}^*\mathbf{v}_i \\ \frac{1}{\sigma_1^2}\sum_{t=1}^T\sum_{i=1}^N d_t\mu_i + N\mathbf{d}'\mathbf{A}^*\boldsymbol{\lambda} + \sum_{i=1}^N \mathbf{d}'\mathbf{A}^*\mathbf{v}_i \\ \frac{T}{\sigma_1^2}\sum_{i=1}^N h_i\mu_i + \frac{1}{\sigma_1^2}\sum_{i=1}^N h_i\sum_{t=1}^T \lambda_t + \frac{1}{\sigma_1^2}\sum_{t=1}^T\sum_{i=1}^N v_{it}h_i \end{bmatrix}\end{aligned}$$

with $\sigma_1^2 = T\sigma_\mu^2 + \sigma_e^2$. Proceeding as in the proof of theorem 1 then obtains the results in (ii) $\blacksquare$

References

Amemiya, T. (1971), 'The estimation of the variances in a variance-components model', *International Economic Review* **12**, 1–13.

Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press.

Anderson, T. W. and Hsiao, C. (1982), 'Formulation and estimation of dynamic models using panel data', *Journal of Econometrics* **18**, 47–82.

Balestra, P. and Nerlove, M. (1966), 'Pooling cross-section and time-series data in the estimation of a dynamic model: The demand for natural gas', *Econometrica* **34**, 585–612.

Baltagi, B. H. and Li, Q. (1991), 'A transformation that will circumvent the problem of autocorrelation in an error component model', *Journal of Econometrics* **48**, 385–393.

Baltagi, B. H. and Li, Q. (1994), 'Estimating errror component models with general MA(q) disturbances', *Econometric Theory* **10**, 396–408.

Dhrymes, P. J. (1984), *Mathematics for Econometrics*, Springer Verlag.

Gourieroux, C. and Monfort, A. (1995), *Statistics and Econometric Models, Volume Two*, Cambridge University Press.

Karlsson, S. and Skoglund, J. (2000), Maximum-likelihood based inference in the two-way random effects model with serially correlated time effects, SSE/EFI Working Paper Series in Economics and Finance 383, Stockholm School of Economics.

King, M. L. (1986), 'Efficient estimation and testing of regressions with a serially correlated error component.', *Journal of Quantitative Economics* **2**, 231–247.

Lillard, L. A. and Willis, R. J. (1978), 'Dynamic aspects of earnings mobility', *Econometrica* **46**, 985–1012.

MaCurdy, T. E. (1982), 'The use of time series processes to model the error structure of earnings in longitudinal data analysis', *Journal of Econometrics* **18**, 83–114.

Magnus, J. R. and Woodland, A. D. (1988), 'On the maximum likelihood estimation of multivariate regression models containing serially correlated error components', *International Economic Review* **29**, 707–725.

Phillips, P. C. and Moon, H. R. (1999), Linear regression limit theory for nonstationary panel data, Cowles Foundation Discussion Papers 1222, Yale University.

Revankar, N. S. (1979), 'Error component models with serially correlated time effects', *Journal of the Indian Statistical Association* **17**, 137–160.

Rudin, W. (1976), *Principles of Mathematical Analysis*, McGraw-Hill.